

# URCHIN: A Reverse Ray Tracer for Astrophysical Applications

Gabriel Altay<sup>1</sup> and Tom Theuns<sup>1,2</sup>

<sup>1</sup>*Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE*

<sup>2</sup>*Department of Physics, University of Antwerp, Campus Groenenborger, Groenenborgerlaan 171, B-2020 Antwerp, Belgium*

Accepted 201? ????. Received 201? ????. in original form 2010 xx

## ABSTRACT

We describe URCHIN, a reverse ray tracing radiative transfer scheme optimised to model self-shielding from the post-reionisation ultraviolet (UV) background in cosmological simulations. The reverse ray tracing strategy provides several benefits over forward ray tracing codes including: (1) the preservation of adaptive density field resolution (2) completely uniform sampling of gas elements by rays; (3) the preservation of galilean invariance; (4) the ability to sample the UV background spectrum with hundreds of frequency bins; and (5) exact preservation of the input UV background spectrum and amplitude in optically thin gas. The implementation described here focuses on Smoothed Particle Hydrodynamics (SPH). However, the method can be applied to any density field representation in which resolution elements admit ray intersection tests and can be associated with optical depths. We characterise the errors in our implementation in stages beginning with comparison to known analytic solutions and ending with a realistic model of the  $z = 3$  cosmological UV background incident onto a suite of spherically symmetric models of gaseous galactic halos.

**Key words:** methods: numerical – radiative transfer – intergalactic medium – quasars: absorption lines – diffuse radiation – ultraviolet: general

## 1 INTRODUCTION

Cosmological gas dynamics simulations (e.g. Cen et al. 1994; Theuns et al. 1998b,a; Hernquist et al. 1996), semi-analytic models (e.g. Bi & Davidsen 1997) and analytic calculations (e.g. Schaye 2001) have enabled us to understand the connection between galaxies, the intergalactic medium (IGM), and the large-scale structure of the Universe. HI Lyman- $\alpha$  absorbers in the spectra of distant quasars are a particularly useful observational probe of this structure. Models indicate that at redshifts  $z = 2 - 4$ , the majority of Lyman- $\alpha$  forest absorption lines (i.e. lines with column densities  $N_{\text{HI}} < 10^{17.2} \text{ cm}^{-2}$ ) arise in filamentary structures in the IGM with the highest column density lines being associated with the circumgalactic medium of galaxies. These lines are produced in systems which are highly ionized by the ultraviolet (UV) background (see Meiksin 2009, for a recent review). An absorber with a column density  $N_{\text{HI}} > 10^{17.2} \text{ cm}^{-2}$  has an optical depth greater than unity for Lyman-Limit photons and is called a Lyman-Limit System (LLS). Above a column density of  $N_{\text{HI}} = 10^{20.3} \text{ cm}^{-2}$ , ‘damping wings’ due to the natural line-broadening of the Lyman- $\alpha$  line are detectable, and the system is called a damped Lyman- $\alpha$  absorber (DLA). These absorbers probe the interface between the IGM and galaxies as well as the interstellar medium (ISM) of the galaxies themselves. Hydrogen begins to self-shield from the UV background in the LLS column density range, and the reduction in ionising flux plays a major role in setting the ionisation state of these absorbers. The URCHIN code described in this paper is designed to model HI self-shielding in the LLS and DLA range.

At present, the largest observational catalogues of self-shielded absorbers are produced through semi-automated searches (e.g. Prochaska & Herbert-Fort 2004; Noterdaeme et al. 2012) of data from the Sloan Digital Sky Survey (SDSS<sup>1</sup>). Current and planned expansions to the SDSS such as the Baryon Oscillation Spectroscopic Survey (BOSS, Schlegel et al. 2007) and BigBOSS (Myers et al. 2012) will increase the amount of available data by a factor of ten. Due to atmospheric absorption of the rest frame Lyman- $\alpha$  transition, ground based surveys for DLAs are limited to redshifts  $z > 1.6$ . Surveys for LLSs require spectral coverage of the Lyman-Limit transition for an accurate determination of  $N_{\text{HI}}$  and are therefore limited to redshifts of  $z > 2.5$  when performed from the ground. The new *Cosmic Origins Spectrograph*<sup>2</sup> on the *Hubble Space Telescope* provides significant capacity to probe lower redshift systems (e.g. Battisti et al. 2012) while the *Advanced Camera for Surveys* and *Wide Field Camera 3* have recently been used to complete a survey for LLSs in the redshift range  $1.0 < z < 2.6$  (O’Meara et al. 2012).

In addition to Lyman series transitions, post-reionisation neutral hydrogen can also be effectively probed using the 21-cm emission line. The most recent determination of the local HI mass function is from the Arecibo Legacy Fast ALFA (ALFALFA) survey (Martin et al. 2010) which will have detected  $\approx 3 \times 10^4$  galaxies in HI 21-cm out to  $z = 0.06$  when it is complete. The Square Kilo-

<sup>1</sup> [www.sdss3.org](http://www.sdss3.org)

<sup>2</sup> [www.stsci.edu/hst/cos](http://www.stsci.edu/hst/cos)

meter Array (SKA<sup>3</sup>) represents the long term future for this type of radio astronomy however construction will not begin for several years. In preparation of this, a host of pathfinding telescopes (ASKAP<sup>4</sup>, MeerKAT<sup>5</sup>, WSRT<sup>6</sup>, EVLA<sup>7</sup>) will soon make 21-cm emission surveys as data rich as their optical and near-infrared counterparts. In addition, pilot surveys for 21-cm absorption in the spectra of radio bright sources have shown potential (e.g. Gupta et al. 2010; Darling et al. 2011) and at least two of these pathfinders (ASKAP and MeerKAT) will also perform large, blind, absorption surveys.

The combined output from these surveys will provide transformative information on the gas content of galaxies and their modes of accretion. It will also generate samples that trace the large scale structure of the Universe with different biases than those of optically selected samples. A prerequisite for making model predictions of hydrogen emission or absorption is the accurate calculation of the distribution of neutral hydrogen (e.g. Duffy et al. 2012; van de Voort et al. 2012). Methods that accomplish this task with a minimum of free parameters will be able to take full advantage of observational data. These issues motivate many of the design choices for URCHIN.

The standard approach of treating the post-reionization UV background in cosmological simulations is to impose a spatially uniform but time varying radiation field, and calculate the H I fraction in the optically thin limit. Pioneering work on modeling self-shielding in gas dynamics simulations was done by Katz et al. (1996) and Haehnelt et al. (1998) by post-processing column density maps. More recent theoretical work has incorporated radiative transfer through 3-D density fields to calculate the attenuation of the UV background in dense gas (Razoumov et al. 2006; Kohler & Gnedin 2007; Pontzen et al. 2008; Altay et al. 2011; McQuinn et al. 2011; Yajima et al. 2011; Fumagalli et al. 2011; Cen 2012; Erkal et al. 2012).

In this paper we present and test URCHIN, a reverse ray tracing scheme designed to calculate self-shielding corrections in the post-reionisation Universe. The code can be applied to any density field representation (e.g. particles, adaptive grids, unstructured meshes) in which the resolution elements can be associated with optical depths and subjected to ray intersection tests. The main benefits of URCHIN are: (1) preservation of the adaptive density field resolution present in many gas dynamics codes; (2) uniform sampling of gas resolution elements with rays; (3) preservation of galilean invariance; (4) high spectral resolution; and (5) preservation of the standard uniform UV background in optically thin gas. The format of this paper is as follows. In §2 we introduce our notation and review some basic physics related to radiative transfer. In §3 we give a general description of our reverse ray tracing approach and place it in context by comparing it to alternative approaches. In §4 we discuss the details of our implementation using smoothed particle hydrodynamics (SPH) density fields. In §5 we present the results of tests meant to validate URCHIN and in §6 we discuss the results, suggest improvements for future versions of the code, and conclude.

## 2 DEFINITIONS AND BASIC PHYSICS

In this section, we define our notation and review some of the relevant physics. All simulations discussed in this work utilize a cubic simulation volume. For brevity, we will refer to any of these simulation volumes as boxes and their six faces as walls. When referring to distances, we will distinguish between proper and comoving measures using the prefixes ‘p’ and ‘c’ (e.g. pkpc, cMpc). In this work, we consider only hydrogen and leave the inclusion of other elements, particularly helium, to future work. For our description of the radiation field, we adopt the notation of Rybicki & Lightman (1986).

The specific intensity,  $I_\nu \equiv dE/dA d\Omega dt d\nu$ , fully characterises a radiation field and is defined as the energy  $dE$  passing through an area element  $dA$  into a solid angle element  $d\Omega$  in time  $dt$  due to photons with frequency between  $\nu$  and  $\nu + d\nu$ . Several useful characterisations of the radiation field can be expressed as integrals over this quantity. Considering photons with frequency  $\nu_{\text{th}} < \nu < q\nu_{\text{th}}$ , and an optically thin medium, we can write the number density of hydrogen ionising photons and the photoionisation rate at the point  $\mathbf{x}$  respectively as:

$$n_\gamma(\mathbf{x}) = \frac{1}{c} \oint \int_{\nu_{\text{th}}}^{q\nu_{\text{th}}} \frac{I_\nu}{h\nu} d\nu d\Omega, \quad (1)$$

$$\Gamma(\mathbf{x}) = \oint \int_{\nu_{\text{th}}}^{q\nu_{\text{th}}} \frac{I_\nu \sigma}{h\nu} d\nu d\Omega \quad (2)$$

where  $h\nu_{\text{th}} = E_{\text{th}}$  and  $\sigma$  are the ionisation energy and photoionisation cross-section of hydrogen. In the case of a medium with finite opacity, the above quantities can be written in terms of the optically thin value of  $I_\nu$  by making the replacement  $I_\nu \rightarrow I_\nu \exp[-\tau(\nu, \Omega)]$  where  $\tau$  is the optical depth between the sources producing  $I_\nu$  and  $\mathbf{x}$ .

The frequency averaged (‘grey’) photoionisation cross section is defined as  $\sigma_{\text{grey}} \equiv \Gamma(cn_\gamma)^{-1}$ . Under the grey approximation, every polychromatic spectrum (characterised by  $n_\gamma$  and  $\Gamma$ ) corresponds to an equivalent monochromatic spectrum with the same  $n_\gamma$  and  $\Gamma$ . The flux of the monochromatic spectrum is fixed by  $n_\gamma$  and the energy of the photons in the monochromatic spectrum is  $E_{\text{grey}} = h\nu_{\text{grey}}$  with  $\nu_{\text{grey}}$  implicitly defined by  $\sigma(\nu_{\text{grey}}) = \sigma_{\text{grey}}$ .

## 3 REVERSE RAY TRACING METHOD

URCHIN is designed to efficiently model the residual neutral hydrogen in the post-reionisation universe. In this section, we describe the algorithms used in URCHIN and conceptual departures from alternative radiative transfer codes. We begin with a brief description of the standard treatment of the UV background in cosmological gas dynamics simulations.

### 3.1 Standard Treatment of Post-Reionisation UV Background

In the post-reionisation Universe, cosmic hydrogen is kept highly ionised by a pervasive UV background (Gunn & Peterson 1965) produced by galaxies and quasars. Quantitatively, the volume averaged neutral fraction  $x = n_{\text{HI}}/n_{\text{H}}$  for redshifts  $z \leq 6$  is on the order of  $10^{-4}$  (Fan et al. 2006; Becker et al. 2007). A rapid transition to higher neutral fractions signals the end of reionisation, evidence for which has recently been observed in the form of a damping wing in the spectrum of a  $z \sim 7$  quasar (Mortlock et al. 2011). At lower redshifts, the UV background determines both the

<sup>3</sup> [www.skatelescope.org](http://www.skatelescope.org)

<sup>4</sup> [www.atnf.csiro.au/projects/askap](http://www.atnf.csiro.au/projects/askap)

<sup>5</sup> [www.ska.ac.za/meerkat](http://www.ska.ac.za/meerkat)

<sup>6</sup> [www.astron.nl/radio-observatory/astronomers/wsrt-astronomers](http://www.astron.nl/radio-observatory/astronomers/wsrt-astronomers)

<sup>7</sup> <http://science.nrao.edu/facilities/evla>

ionisation state and temperature of gas in the IGM and sets the rate at which denser gas can cool, accrete onto small galaxies, and form stars (Efsthathiou 1992; Okamoto et al. 2008).

The most widely used treatment of the post-reionisation UV background in cosmological simulations, is based on three approximations: (1) optically thin gas; (2) a spatially uniform UV background; and (3) photo/collisional ionisation equilibrium. These approximations are valid for the majority of cosmic gas, i.e. for highly ionised hydrogen; however, they break down for the majority of gas observable in HI surveys. The approximation of optically thin gas begins to break down in absorption systems that probe accretion (LLSs) from the IGM onto galaxies and completely fails in regions of significant self-shielding (DLAs) where the strongest HI signals arise. This approximation is the most important to remedy for HI surveys. The second approximation involves disregarding large scale gradients in the UV background as well as point sources. These fluctuations can have an effect on absorber statistics (e.g. Rahmati et al. 2013; Yajima et al. 2011; Croft 2004) but the magnitude is not as large as that due to self-shielding. A notable exception is absorbers in the proximity zones of bright sources (Schaye 2006). The third approximation involving photo/collisional equilibrium likely holds for dense gas where HI is self-shielded and recombination times are short compared to UV background variability, but will break down near variable sources and in less dense gas. Relaxing these approximations is an efficient way to improve model predictions for HI surveys. In this release of URCHIN we focus on self-shielding.

### 3.2 Reverse Ray Tracing - Motivation

Numerical techniques for continuum radiative transfer have been developed based around ray tracing methods, (e.g. Nakamoto et al. 2001; Maselli et al. 2003; Razoumov & Cardall 2005; Susa 2006; Whalen & Norman 2006; Altay et al. 2008), the closely related method of characteristics (e.g. Mellema et al. 2006; Rijkhorst et al. 2006), angular moments of the radiative transfer equation (e.g. Gnedin & Abel 2001; Aubert & Teyssier 2008; Petkova & Springel 2009; Finlator et al. 2009), and transport on unstructured meshes (e.g. Pawlik & Schaye 2008, 2011; Paardekooper et al. 2010). A detailed comparison between many codes currently in use is documented in the Cosmological Radiative Transfer Comparison Project (Iliev et al. 2006, 2009).

The above methods are all based on following radiation from its source to the point where it is absorbed or scattered. When dealing with the post-reionisation UV background, the goal is to build up a radiation field that is known from both theoretical and observation studies to be mostly uniform. In these methods, the UV background field at a given point is the sum of radiation that has been transported from all sources being considered. However, in many methods this is true only in a statistical sense. For example, in Monte Carlo ray tracers such as SPHray, resolution elements are updated whenever a ray intersects them. The ray only carries information about the flux from one source, but if enough rays are traced, each resolution element will be updated by the rays from many sources. In URCHIN we attempt a different type of solution to this problem. As opposed to building up a mostly uniform UV background from multiple sources, we begin with the standard approximations described above (optically thin gas, uniform field, photo-collisional equilibrium) and then calculate deviations from it. For the majority of cosmic gas, the optically thin photoionisation rate from this uniform field,  $\Gamma^{\text{thin}}$ , is in fact a good approximation. The current version of URCHIN relaxes the optically

thin approximation from the standard treatment by attenuating this uniform radiation field in denser regions that begin to self-shield. The fraction of gas (by mass or volume) where this correction is necessary is guaranteed to be small by the nature of the problem allowing us to concentrate the available computational resources where they are needed. In future versions of the code we will relax the second and third approximations in the standard treatment by adding large scale gradients, proximity regions, and considering non-equilibrium effects. Currently URCHIN operates on static density fields as a post-processing step, but in principle could be coupled to existing gas dynamics codes in a straight-forward way (e.g. in a framework such as described in Portegies Zwart et al. 2009).

### 3.3 Reverse Ray Tracing - Algorithm

Consider a density field discretised into resolution elements labeled  $i = 1, \dots, N$ . We will refer to these resolution elements as particles; however, our reverse ray tracing technique can be applied to any density field representation in which the resolution elements can be associated with optical depths and subjected to ray intersection tests. From each particle, we cast  $N_{\text{ray}}$  ray segments out to a distance  $l_{\text{ray}}$ , in directions that uniformly cover the solid angle around the particle. We use the HEALPix algorithm (Górski et al. 2005) to determine ray directions. We then calculate the optical depth  $\tau_k$  along each of these ray segments and sum over all rays to obtain an estimate of the self-shielded photoionisation rate,  $\Gamma^{\text{shld}} \leq \Gamma^{\text{thin}}$ , at the location of each particle,

$$\begin{aligned} \frac{\Gamma^{\text{shld}}}{\Gamma^{\text{thin}}} &\equiv \exp(-\tau_{\text{eff}}) \\ &= \frac{1}{\Gamma^{\text{thin}}} \frac{4\pi}{N_{\text{ray}}} \sum_{k=1}^{N_{\text{ray}}} \int_{\nu_{\text{th}}}^{q\nu_{\text{th}}} \frac{I_{\nu}\sigma}{h\nu} \exp(-\tau_k) d\nu. \end{aligned} \quad (3)$$

For most particles the optically thin approximation is very good and the effective optical depth will be small,  $\tau_{\text{eff}} \ll 1$ , along all  $N_{\text{ray}}$  directions. However, for the small fraction of particles that dominate the HI abundance, the effective optical depth will be large ( $\tau_{\text{eff}} \gg 1$ ). We then use  $\Gamma^{\text{shld}}$  in an analytic solution for the equilibrium neutral fraction  $x$  (Eq. A9 in the Appendix) to update the ionisation state of the particle. An altered neutral fraction for one particle leads to altered optical depths along ray segments that pass through it, and hence we iterate this procedure until the neutral fraction converges for all particles. This typically happens in tens of iterations with those few particles at the threshold between optically thin and optically thick converging last.  $l_{\text{ray}}$  and  $N_{\text{ray}}$  are numerical parameters of the scheme, and we have found that in fully cosmological runs at  $z \sim 3$ , values of  $l_{\text{ray}} = 100$  pkpc and  $N_{\text{ray}} = 12$  lead to converged column density distribution functions.

### 3.4 Reverse Ray Tracing - Advantages

Reverse ray tracing presents some important advantages in the post-reionization regime, related to the dramatic differences in the character of the radiation field during and after reionization. In the following sub-sections, we focus on questions that typically occur in ray tracing schemes attempting to model the post-reionization cosmological UV background: 1) Where should rays originate and terminate? 2) How should rays be traced to uniformly sample the gas elements? 3) How much spectral resolution can be attained? and 4) What optimizations can be applied? We will show that answers to

these questions are simpler in reverse ray tracing schemes than in forward versions.

### 3.4.1 Where Should Rays Originate?

In forward ray tracing schemes, the UV background at a given point in space is the result of transporting photons along rays originating at sources of radiation. These sources can be divided into two categories: those inside the box (internal sources) and those outside the box (external sources). Modelling the UV background from first principles using only internal sources imposes stringent conditions on the size of the box. A reasonable absolute minimum scale is one mean free path at the Lyman limit,  $\lambda_{912}^{\text{mfp}}$ . Prochaska et al. (2009) find that this mean free path depends on redshift  $z$  approximately as  $\lambda(z) \approx [(48 \pm 2.1) - (38.0 \pm 5.3)(z - 3.6)] h_{72}^{-1}$  pMpc for  $3.6 < z < 4.3$ . At  $z = 3.6$  this is larger than the majority of cosmological gas dynamics simulations able to resolve any galaxy formation processes and  $\lambda_{912}^{\text{mfp}}$  becomes larger at lower redshifts. For this reason, forward ray tracing schemes inevitably must rely on tracing many rays from external sources in addition to internal ones.

The contribution from external sources is usually modeled by casting rays inward from the walls. These rays are then followed until a given fraction of their initial photon content is absorbed. The most straight forward method of choosing ray origins is a pseudo-random sampling of points on the walls. The appropriate flux is usually determined by requiring that the photoionisation rate in optically thin gas match an input value. This method has two drawbacks. The first, discussed in the next sub-section, is the difficulty in uniformly sampling an adaptive density field. The second is the artificial gradient in photoionisation rate that will develop between regions near the centre of the box, where the background rays will be most attenuated, and regions near the walls, where the background will be at its optically thin strength. This is what we term the galilean invariance problem. This adds ambiguity to any calibration of flux from the walls and makes the calibration dependent on the box size. In general, tracing rays from the walls leads unavoidably to a loss of symmetry in the UV background and difficulty in flux calibration. These problems can be alleviated somewhat by tracing rays from randomly selected under-dense regions within the box (Maselli & Ferrara 2005), but not eliminated entirely. Some authors (e.g. McQuinn et al. 2011) have circumvented this problem by excising small regions around halos and ray tracing them individually however this neglects filamentary gas between halos.

In a reverse ray tracing method, these problems are not present. All symmetries of the UV background are maintained and the photoionisation rate at a given particle position is independent of any properties of the box. In addition, the intensity of the UV background in optically thin gas is equal to its value in the standard approximation (optically thin gas, uniform field, photo-collisional equilibrium) by construction. This conveniently allows for direct comparisons between results with and without radiative transfer with no need for calibrating fluxes.

### 3.4.2 The Uniform Sampling Problem

Modern gas dynamical simulations discretise density fields with adaptive resolution elements. This typically leads to better spatial resolution (i.e. smaller resolution elements) in regions of higher gas density. A radiative transfer approach which relies on rays uniformly cast from the simulation walls leads to poor sampling of the

resolution elements. Specifically, the overdense self-shielded regions where radiative transfer is most important are undersampled, and the underdense optically thin regions where radiative transfer is unnecessary are oversampled. A scheme in which rays split and merge (Wise & Abel 2011) can maintain a constant number of rays intersecting each resolution element but comes at the cost of increased overhead. On the other hand, the reverse ray-tracing algorithm samples each particle with the same number of rays by construction. This helps focus computational power where it is needed.

### 3.4.3 Spectral Resolution

In a forward ray tracing scheme, the radiation field at a particle position is the summation of contributions from different rays. For monochromatic spectra, all attenuation information can be characterised by the number of photons in a ray. To accurately handle multi-frequency spectra, one must substantially increase the number of monochromatic rays traced or carry spectral information along with each ray. Similarly for moment methods, each frequency group must be treated separately by the solver. Low fidelity sampling of the UV background spectrum can lead to errors up to an order of magnitude in HII region, see e.g. (Mirocha et al. 2012) and Figure 3 below.

In reverse ray tracing, the rays simply sample the optical depth along a particular direction. Subsequently, arbitrarily complex input spectra can be numerically integrated over frequency to determine a shielded photoionisation rate, and update the neutral fraction of a particle based on full knowledge of the amplitude and spectral shape of the local radiation field. When using 100 frequency bins in URCHIN, the numerical integration over the UV background spectrum consumes a negligible fraction of the computing time. This allows for studies of spectral hardness and eases the inclusion of sharp features such as recombination lines into the UV background spectrum. As an added bonus, knowledge of the optical depth in all directions around a particle during the update allows one to choose the appropriate ‘case A’ (recombinations to all levels) or ‘case B’ (recombinations to all but the ground state) recombination rates in the on-the-spot approximation.

### 3.4.4 Optimisations

The reverse ray tracing approach also lends itself to some important optimisations that are not easy to implement in forward ray tracing techniques. The most expedient is simply avoiding radiative transfer where it is unnecessary. As discussed above, particles that are not self-shielded and not in proximity zones are treated correctly in the uniform and optically thin limit. When looping over the particles in reverse ray tracing schemes, those that satisfy this criterion can be skipped. Identifying these particles is implementation dependent; however, in cosmological simulations, the majority of post-reionization gas does not need to participate in radiative transfer. We will discuss our technique for identifying these particles in the next section.

There is also a sense of ray locality and ray independence inherent in the method that can be useful in optimisations. Any sub-volume in a box can be treated independent of any other as long as a buffer of  $l_{\text{ray}}$  is included around the sub-volumes. In addition, each ray can be made independent during a single iteration by waiting to update the neutral fractions of particles until shielded photoionisation rates have been calculated for all of them. This strategy may increase the number of iterations necessary for convergence but would make each iteration much faster.

#### 4 IMPLEMENTATION DETAILS

URCHIN applies a specified radiation field to a given density field to determine the properties of self-shielded regions. Spatially uniform models of the cosmological UV background which provide  $I_\nu(z)$ , are publically available (for example Haardt & Madau 2012)<sup>8</sup>. The main parameters in URCHIN are the frequency range of ionising photons included in the calculation, the number of rays casts per particle,  $N_{\text{ray}}$ , their proper length,  $l_{\text{ray}}$ , the criterion for choosing between type A and type B recombination rates, and the condition for convergence  $\delta x_{\text{tol}}$ . The choice of spectrum and frequency range determines the optically thin photoionisation rate,  $\Gamma_{\text{thin}}$ , from Eq. (2). Our default runs use  $\nu_{\text{th}} < \nu < 4\nu_{\text{th}}$ ,  $N_{\text{ray}} = 12$  and  $l_{\text{ray}} = 100$  pkpc, which give converged numerical results at redshift  $z = 3$ . In addition, we use case B recombination rates for particles with  $\tau_{\text{eff}} \geq 1$ .

Initially the neutral fraction  $x$  of each particle is set to its optically thin value,  $x_{\text{thin}} = n_{\text{HI}}^{\text{thin}}/n_{\text{H}}$ . This allows for the calculation of the HI column density along each ray cast from a particle. Next we calculate a shielded photoionisation rate,  $\Gamma^{\text{shld}}$ , for each particle and an effective optical depth,  $\tau_{\text{eff}}$  using Eq. (3). Particles with  $\tau_{\text{eff}} < \tau_{\text{eff}}^{\text{skip}}$  maintain  $x = x_{\text{thin}}$  for all subsequent iterations while all other particles are updated each iteration. We continue to loop over particles which initially had  $\tau_{\text{eff}} > \tau_{\text{eff}}^{\text{skip}}$  until convergence in the neutral fraction,  $|\delta x|/x < \delta x_{\text{tol}}$ . Typical values for these parameters are  $\tau_{\text{eff}}^{\text{skip}} = 10^{-4}$ ,  $\delta x_{\text{tol}} = 10^{-3}$ ,  $l_{\text{ray}} = 100$  pkpc and  $N_{\text{ray}} = 12$ . For these choices we find that more than 99 percent of particles have converged neutral fractions after five iterations in a cosmological run, with the remaining 1 percent requiring tens of iterations. In our current implementation, we update the neutral fraction of each particle as soon as its self-shielded photoionisation rate has been computed. This update scheme works well provided the list of particles is traversed from high to low density, however, there is no fundamental restriction on when the particles should be updated. Implementations which calculate  $\Gamma^{\text{shld}}$  for each particle before performing any updates of the neutral fractions are independent of the order in which the particles are looped over, which could be useful in parallel strategies.

##### 4.1 Reverse Ray Tracing in SPH

Although URCHIN can be applied to many types of density field discretisations (particles, grids, unstructured meshes), in this section we discuss our implementation in Smoothed Particle Hydrodynamics (SPH, Gingold & Monaghan 1977; Lucy 1977).

###### 4.1.1 Column Densities in SPH

Our calculation of SPH column densities makes use of several improvements over the algorithm described in Altay et al. (2008). In the SPH formalism the number density of HI at any location  $\mathbf{r}$  can be calculated using the scatter approach as

$$n_{\text{HI}}(\mathbf{r}) = \sum_i \frac{m_i x_i}{m_p} W(q_i), \quad (4)$$

where  $m_p$  is the mass of a proton,  $m_i$  is the mass of particle  $i$  in hydrogen,  $x_i$  is its neutral fraction,  $q_i \equiv |\mathbf{r} - \mathbf{r}_i|/h_i$  is the distance between the particle and the point  $\mathbf{r}$  in units of the particle's smoothing length  $h_i$ , and  $W$  is the SPH smoothing kernel.

The column density through an SPH distribution along a path  $\mathbf{r}(l)$  parameterised by  $0 < l < L$  can then be written as:

$$\begin{aligned} N_{\text{HI}} &= \int_0^L n_{\text{HI}}(\mathbf{r}) dl = \int_0^L \sum_i \frac{m_i x_i}{m_p} W(r_{il}, h_i) dl \\ &= \sum_i \frac{m_i x_i}{m_p} \int_0^L W(q_{il}) dl, \end{aligned} \quad (5)$$

where the summation is over particles with smoothing volumes intersected by the path  $\mathbf{r}(l)$  and the subscripts in the variable  $q_{il}$  indicate that it is a function of the summation index  $i$  and distance along the path  $l$ . In this way, the calculation of optical depths is reduced to calculating which particles are intersected by a ray and line integrals through the smoothing kernel  $W$ .

The Gaussian function has many properties that make it a natural choice for the smoothing kernel, however, its lack of compact support leads to an impractical sum over all particles in the simulation volume. To remedy this, many SPH codes make use of spline functions with an approximately Gaussian shape, such as the  $M_4$  cubic spline (Monaghan & Lattanzio 1985),

$$M_4(q) = \frac{8}{\pi h^3} \begin{cases} 1 - 6q^2 + 6q^3 & \text{for } 0 \leq q \leq \frac{1}{2} \\ 2(1 - q)^3 & \text{for } \frac{1}{2} < q \leq 1 \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

In URCHIN we use a truncated Gaussian kernel that allows us to obtain line integrals at a given impact parameter analytically. A normalised Gaussian function centred at the origin is given by

$$G(r, \sigma^2) = \frac{\exp(-A^2 r^2)}{(2\pi\sigma^2)^{3/2}}, \quad (7)$$

where  $A^2 \equiv (2\sigma^2)^{-1}$  and  $\sigma^2$  is the variance. We truncate  $G$  at  $r = h$  and further demand that  $G(0, \sigma^2) = M_4(0, h)$  to obtain the truncated Gaussian kernel

$$G_t(r, \sigma) = \mathcal{N} \begin{cases} \exp(-A^2 r^2) & \text{for } r \leq h \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where

$$\begin{aligned} \sigma^2 &= \frac{h^2}{8\pi^{1/3}} \\ A^2 &= \frac{4\pi^{1/3}}{h^2} \\ \mathcal{N} &= \frac{8}{\pi h^3} \left[ \text{erf}(t) - \frac{2t \exp(-t^2)}{\sqrt{\pi}} \right]^{-1} \\ t &= 2\pi^{1/6}, \end{aligned} \quad (9)$$

and the error function is defined as,

$$\text{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t \exp(-t^2) dt. \quad (10)$$

The column density through such a Gaussian kernel at impact parameter  $b$ , between the limits  $z_1$  and  $z_2$  is then

$$I(h, b, z_1, z_2) = \frac{\mathcal{N} \exp(-A^2 b^2)}{A} [\text{erf}(Az_2) - \text{erf}(Az_1)]. \quad (11)$$

We use this kernel to calculate the HI column densities along each ray using Eq. (5). Incidentally, the fact that the Gaussian kernel can be decomposed into three 1-D functions makes it useful for smoothing SPH particles onto 2-D or 3-D grids.

<sup>8</sup> <http://www.icolick.org/~pmadau/CUBA>

4.1.2 *Self-contribution to self-shielding in SPH*

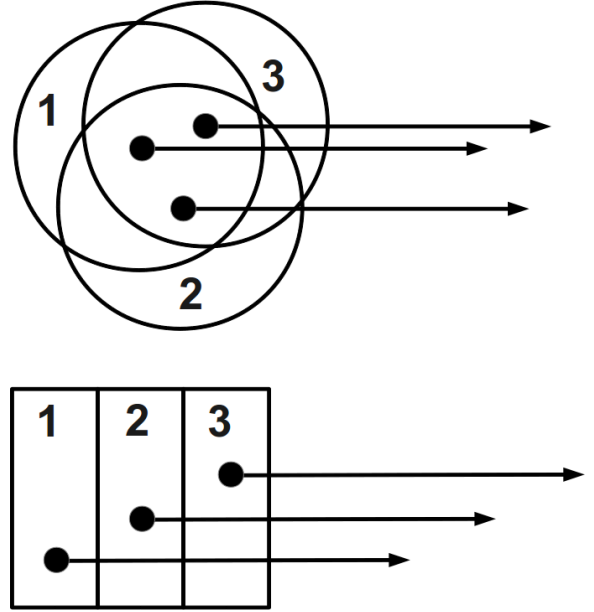
When passing from a continuous density field representation to a discrete one, Eq. (A9) for the equilibrium neutral fraction at the point  $\mathbf{r}$ ,  $x = x(\mathbf{r}, \Gamma, n_{\text{H}}, T, y)$  becomes partially implicit: in the discrete case,  $\Gamma$  has a dependence on  $x$  through the particle's own contribution to the optical depth along a ray. This is a generic problem of discretisation and has been discussed previously in Abel et al. (1999) and Mellema et al. (2006).

However when the density field is represented with SPH particles, further complications arise due to the weighted sum over neighbours. For example the HI density at location  $\mathbf{r}$  depends on the neutral fractions of all particles that appear in the sum of Eq. (4). This makes a straightforward scheme that updates neutral fractions based on  $\Gamma^{\text{shld}}$  calculated at the center of an SPH particle unstable. The reason for this instability is illustrated in the top part of Figure 1. Consider three SPH particles which are initially optically thin but are located in a region that will eventually become self-shield. Rays traced outward (in any direction) from each particle will probe optical depth contributed by all three. This will in turn cause the neutral fraction of each particle to be increased. In the next iteration, each ray will find increased optical depth and each neutral fraction will increase again. This process will continue and eventually spread to adjacent particles causing unphysical growth of neutral regions. The cause of this instability is the multi-valued relationship between points in space and resolution elements in SPH. As a counterexample, consider the uniform grid density field represented in the bottom part of Figure 1. This discretisation has a single-valued relationship between points in space and resolution elements. In this case, changes in the neutral fraction of each element do not necessarily change the optical depth encountered by each ray. For example, an increase in the neutral fraction of element 1 has no effect on the optical depth encountered by rays traced from elements 2 or 3. This prevents the instability from occurring.

We resolved this numerical artefact in SPH as follows. Particles intersected by each ray are split into two distinct groups labeled *near* and *far*, with  $\tau = \tau^{\text{near}} + \tau^{\text{far}}$ . Near particles of particle  $i$  are those that contribute to  $i$ 's (neutral) density in Eq. (4), *i.e.* the particle's neighbours. All other particles that contribute to the sum in Eq. (5) are labelled 'far'. The numerical instability only involves near particles, therefore we can treat the *far* particles with the algorithm described in §3. We model the near particles as a uniform slab with the same temperature  $T$  and density  $n_{\text{H}}$  as particle  $i$ . The thickness of the slab is quantified by calculating the total hydrogen optical depth of the *near* particles at the Lyman Limit defined as  $\tau_{\text{H}}^{\text{near}} \equiv N_{\text{H}}^{\text{near}} \sigma_{\text{th}}$ . The column density  $N_{\text{H}}^{\text{near}}$  is calculated as in Eq. (5) except all of the neutral fractions are set to unity.

The radiation incident onto the slab is determined by the user supplied spectrum and the optical depth through the *far* particles,  $I_{\nu} \exp(-\tau^{\text{far}})$ . For each ray  $k = 1, \dots, N_{\text{ray}}$  this incoming flux provides a photoionisation rate at the surface of the slab of  $\Gamma_k^{\text{far}}$ . We then solve for the ionisation structure in the slab and associate the photoionisation rate at the bottom,  $\Gamma_k^{\text{shld}}$ , with the contribution from ray  $k$  to the total shielded photoionisation rate  $\Gamma^{\text{shld}}$  of particle  $i$ . In the case of monochromatic radiation, there is an analytic solution for photoionisation rate as a function of depth into the slab which can be used to determine  $\Gamma_k^{\text{shld}}$  (see Appendix A4). For polychromatic spectra we tabulate the solution as a function of four variables,  $\{\Gamma_k^{\text{far}}/n_{\text{H}}, \tau_k^{\text{far}}, T, \tau_{\text{H},k}^{\text{near}}\}$ <sup>9</sup>.

The first variable,  $\Gamma_k^{\text{far}}/n_{\text{H}}$ , is the ratio of the amplitude of the



**Figure 1.** An illustration of two density field discretisation methods. The top panel shows a multi-valued relationship between points in space and resolution elements (SPH). The bottom panel shows a single valued relationship between points in space and resolution elements (rectilinear). We show a single ray of length  $l_{\text{ray}}$  traced from each resolution element. In the SPH case, a change in the neutral fraction of any element changes the optical depth calculated along all three rays. In the rectilinear case, changing the neutral fraction of element 1 (for example) leaves the optical depth calculated along the rays from elements 2 and 3 unchanged.

incident radiation field over the density of the slab, the second,  $\tau^{\text{far}}$ , determines the incident spectrum (*i.e.* how much the user supplied spectrum has been hardened before entering the slab), the third,  $T$ , determines recombination and collisional ionization rates, and the fourth,  $\tau_{\text{H}}^{\text{near}}$  is the thickness of the slab. In both the analytic and the lookup table case, we label this solution  $\mathcal{G}$  and note that the values of  $\tau^{\text{far}}$ ,  $\Gamma_k^{\text{far}}$ , and  $\tau_{\text{H}}^{\text{near}}$  are different for each of the  $N_{\text{ray}}$  rays traced from particle  $i$ . In summary, the total shielded photoionisation rate  $\Gamma^{\text{shld}}$  for particle  $i$  is computed in the following way. We trace rays  $k = 1, \dots, N_{\text{ray}}$  and calculate,

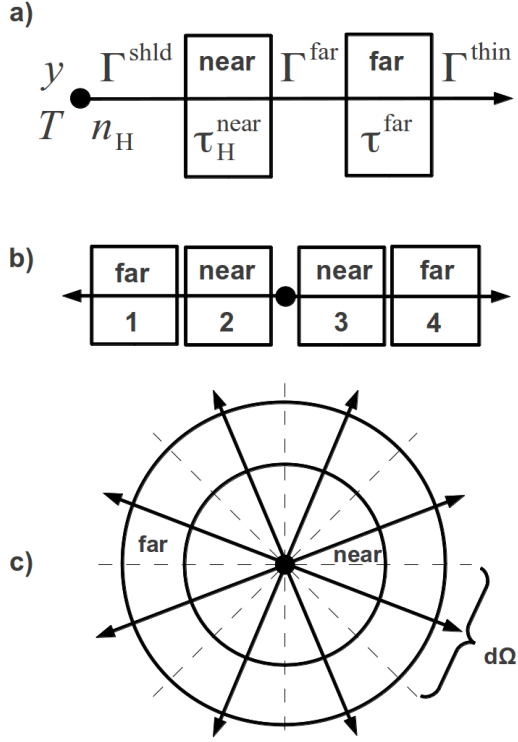
$$\Gamma_k^{\text{far}} = \int_{\nu_{\text{th}}}^{q\nu_{\text{th}}} \frac{I_{\nu} \sigma}{h\nu} e^{-\tau_k^{\text{far}}} d\nu \quad (12)$$

$$\Gamma_k^{\text{shld}} = \mathcal{G} \left( \frac{\Gamma_k^{\text{far}}}{n_{\text{H}}}, T, \tau_k^{\text{far}}, \tau_{\text{H},k}^{\text{near}} \right) \quad (13)$$

$$\Gamma^{\text{shld}} = \sum_{k=1}^{N_{\text{ray}}} \frac{4\pi}{N_{\text{ray}}} \Gamma_k^{\text{shld}}. \quad (14)$$

This whole process is illustrated in Fig. 2. The determination of  $\Gamma^{\text{shld}}$  has no dependence on the neutral fraction of particle  $i$  or any of its neighbors and therefore the solution is numerically stable. However, it requires extra computational effort to evaluate the function  $\mathcal{G}$  and introduces errors due to the lack of interaction between different *near* slabs. In the following section we will quantify these errors.

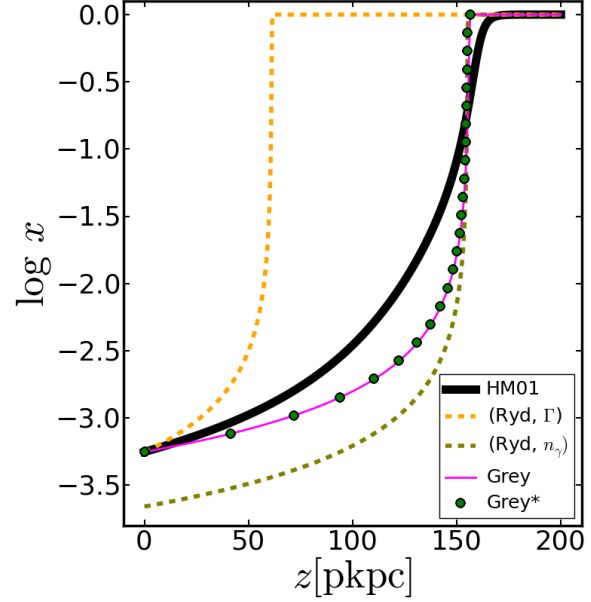
<sup>9</sup> The optical depth  $\tau^{\text{far}} = N_{\text{H}}^{\text{far}} \sigma$  is a function of frequency and not a scalar like the other parameters, however, the full shape of  $\tau^{\text{far}}(\nu)$  is determined by a single evaluation, for example  $\tau^{\text{far}}(\nu_{\text{th}})$ .



**Figure 2.** An illustration of our reverse ray tracing technique in SPH density fields. Panel a) shows a typical situation in Test 2 with a single ray being traced from a particle with density  $n_H$  and temperature  $T$ . The *far* slab on the right indicates the optical depth  $\tau_H^{\text{far}}$  due to *far* particles which attenuate the optically thin radiation field and determine  $\Gamma^{\text{far}}$  (Eq. 12). This, in turn, determines the radiation incident on the slab used to model the *near* particles. The density and temperature of the *near* slab are determined by the particle being updated ( $n_H, T$ ) but its thickness  $\tau_H^{\text{near}} = N_H^{\text{near}} \sigma_{\text{th}}$  is fixed by all *near* particles. The solution  $\mathcal{G}$  is then used to determine the photoionisation rate  $\Gamma^{\text{shld}}$  at the bottom of the slab independent of the neutral fractions of any of the *near* particles (Eq. 13). The variables  $T, n_H, y$ , and  $\Gamma^{\text{shld}}$  are then used in an analytic solution (Eq. A9) to determine the updated  $x$  for the particle. Panel b) shows a typical situation in Test 3, in which two rays are traced from each particle and a  $\Gamma^{\text{shld}}$  is determined for each ray. These are then combined to find the total shielded photoionisation rate for the particle (Eq. 14). In this case, radiation that should be incident on slab 2 from the direction of slab 3 is not accounted for by the solution  $\mathcal{G}$  and vice versa. Panel c) shows a 2-D cartoon of the situations in Test 4 in which rays are traced in all directions. A component of the errors in these tests is due to an extension of the problem in using  $\mathcal{G}$  described for panel b).

## 5 TESTS AND VERIFICATION

In this section, we present several tests performed in order to validate URCHIN. We begin with simple test cases with a known analytic solution, and end with more realistic tests that involve gaseous galactic halos drawn from a cosmological simulation. For those tests in which an analytic solution is not available, we compare the URCHIN results to those of a straight-forward numerical solver which we will call  $\mathcal{NS}$  to distinguish it from URCHIN. After we have verified  $\mathcal{NS}$  against analytic solutions, we will simply refer to  $\mathcal{NS}$  as the analytic solution.



**Figure 3.** Test 1: Equilibrium neutral fraction  $x$  as a function of depth  $z$  in the case of plane-parallel radiation incident onto a slab of hydrogen with constant and uniform density  $n_H = 1.5 \times 10^{-3} \text{cm}^{-3}$  and temperature  $T = 10^4 \text{K}$ . The *thick black line* is the solution when using the spectrum of Haardt & Madau 2001 restricted to photons with energy  $1 < h\nu/h\nu_{\text{th}} < 4$  ( $I_\nu^{\text{HM}}$ ). The *pink line* is the  $\mathcal{NS}$  result for the case of monochromatic photons with energy  $h\nu = 19.2 \text{eV}$  (the grey approximation of  $I_\nu^{\text{HM}}$ ); green dots are the corresponding analytic result discussed in the Appendix. The excellent agreement demonstrates the accuracy of  $\mathcal{NS}$ . By construction, the grey approximation reproduces both the number density of photons,  $n_\gamma$ , and the photoionisation rate,  $\Gamma$ , of  $I_\nu^{\text{HM}}$  in the optically thin limit. The *dashed yellow* and *dashed olive* curves represent cases in which monochromatic photons with  $h\nu = 13.6 \text{eV}$  were incident and the flux was normalised to reproduce either  $\Gamma$  or  $n_\gamma$ , respectively. These monochromatic approximations to HM01 give inaccurate results either in the optically thin limit, and/or in the position of the ionisation front, illustrating the need to properly sample the input spectrum.

### 5.1 Test 1: Analytic Slab Solution

This test involves plane-parallel radiation with flux  $F$  incident from one side onto a slab of hydrogen gas of thickness  $L_{\text{slab}}$ , uniform density  $n_H$ , and fixed uniform temperature  $T_{\text{slab}}$ . The surface of the slab is coincident with the  $z = 0$  plane and the bulk extends in the  $z > 0$  direction. In the case of monochromatic radiation, this problem has an analytic solution which we derive in Appendix A4. The purpose of this test is to verify  $\mathcal{NS}$  and to illustrate the dependence of the solution on the assumed spectrum of the incoming radiation.

The equilibrium neutral fraction as a function of depth,  $x(z)$ , is obtained with  $\mathcal{NS}$  by dividing the slab into many thin slices perpendicular to the  $z$ -axis. Starting from  $z = 0$  and working downwards, we solve for  $x$  in one slice at a time using the following algorithm: 1) determine the HI optical depth,  $\tau = \int_0^z x n_H dz$ , above the current slice; 2) calculate an attenuated photoionisation rate in the slice,  $\Gamma = \int_{\nu_{\text{th}}}^\infty F \sigma \exp(-\tau) d\nu$ ; 3) determine  $x$  in the slice by plugging  $\Gamma$  into an analytic solution (Eq. A9). To avoid errors due to finite slice width, we choose the number of slices such that each has a total hydrogen optical depth  $\tau_H = N_H \sigma_{\text{th}}$  below unity. This guarantees they will always be optically thin when considered individually. The numerical values for the parameters of this test are,



$L_{\text{slab}} = 200$  pkpc,  $n_{\text{H}} = 1.5 \times 10^{-3} \text{cm}^{-3}$  (500 times the cosmic mean  $n_{\text{H}}$  at redshift 3), and  $T_{\text{slab}} = 10^4$  K. We compare the analytical solution to that obtained by  $\mathcal{N}\mathcal{S}$  in Fig. 3 (pink line labelled grey versus green symbols): they agree very well.

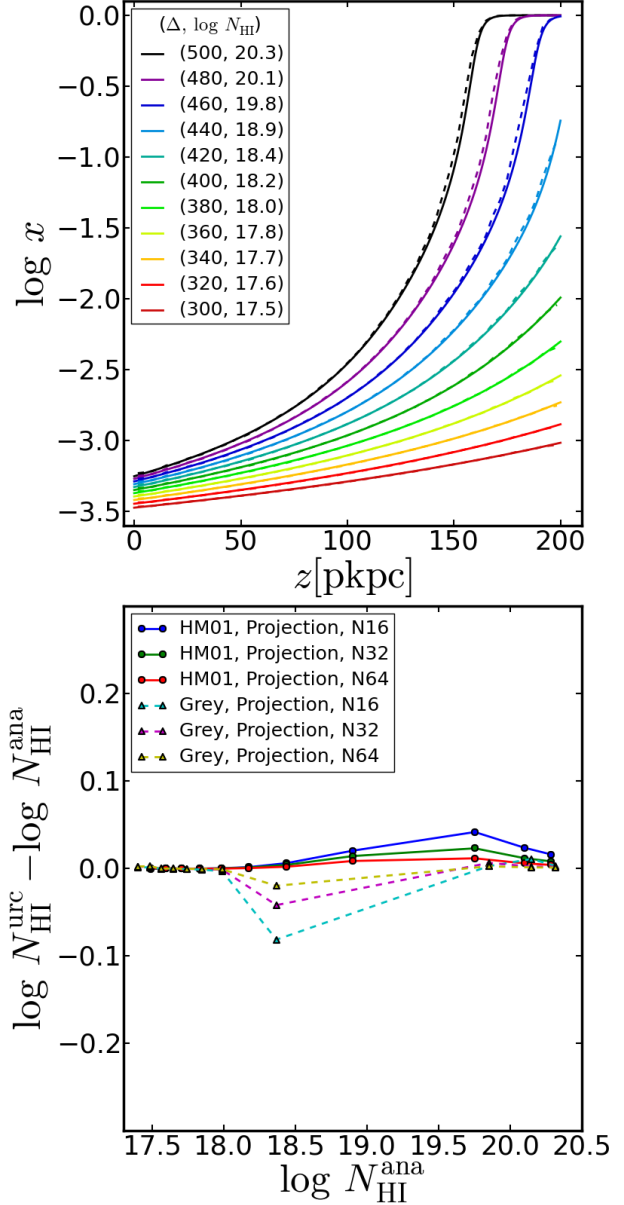
An advantage of the reverse ray tracing approach is the high fidelity with which the spectrum of the ionising background can be sampled. This is non-trivial in forward ray tracing schemes and many rely on a simplified spectrum (see discussion in Mirocha et al. 2012). We can use  $\mathcal{N}\mathcal{S}$  to quantify the accuracy of such approximations. The slab parameters were chosen in order to produce a fully neutral region on the far side of the slab and a total column density  $N_{\text{HI}} = 10^{20.3} \text{cm}^{-2}$  when the redshift  $z = 3$  Haardt & Madau 2001 (HM01) UV background is incident. In what follows, we will refer to the specific intensity of the HM01 UV background at redshift 3 between the frequencies  $\nu_{\text{th}}$  and  $4\nu_{\text{th}}$  as  $I_{\nu}^{\text{HM}}$ . In Figure 3 we compare HM01 with three different monochromatic approximations: *i*) 1 Rydberg photons with a flux that produces the same (optically thin) density of ionising photons as  $I_{\nu}^{\text{HM}}$ , *ii*) 1 Rydberg photons with a flux that produces the same (optically thin) photoionisation rate as  $I_{\nu}^{\text{HM}}$ , and *iii*) the grey approximation of  $I_{\nu}^{\text{HM}}$  which reproduces both the number density of photons and photoionisation rate of  $I_{\nu}^{\text{HM}}$ . Photons in the grey approximation have an energy of 19.2 eV.

In all cases the monochromatic solutions transition to fully neutral more abruptly than the true HM01 solution. This is due to the photons having a fixed photoionisation cross-section. The spread in frequencies in the HM01 spectrum smooths the transition from highly ionised to fully neutral. In the worst case, this can cause an error of several orders of magnitude in the neutral fraction. The monochromatic spectrum normalised to the same photo-ionisation rate (yellow line) recovers the correct neutral fraction in the optically thin region but underestimates the depth of the ionised region by almost 100 pkpc. The monochromatic spectrum with the same number density of ionising photons (olive line) underestimates the neutral fraction in the optically thin region by  $\sim 0.5$  dex, but recovers the location of the ionisation front to better than 10 kpc. The grey approximation (pink line) recovers the ionised fraction in the optically thin region and the location of the ionisation front, but still has a maximum error of  $\sim 0.75$  dex.

The dependence of  $x(z)$  on the spectrum of radiation illustrates the need for accurate multi-frequency treatments. An advantage of URCHIN is that it can treat arbitrary spectral shapes accurately with approximately 100 frequency bins. In addition, the full attenuated ionising spectrum at a particle position is known when neutral fractions are calculated. For completeness we examine several other frequently used approximations to the HM01 spectrum in Appendix A2, and various fitting formula used to approximate the hydrogen photoionisation cross section in Appendix A1.

## 5.2 Test 2: Uniform Slab - Radiation Incident from One Side

In this test we compare URCHIN solutions to those of  $\mathcal{N}\mathcal{S}$ . To this end, we create a set of uniform slabs with the same geometry as in Test 1, but vary the volume densities  $n_{\text{H}}$  such that the projected HI column densities through the slabs cover the range  $17.5 < N_{\text{HI}}/\text{cm}^{-2} < 20.3$  (i.e. the range over which self-shielding becomes important). To create SPH realizations of the uniform density fields we generate glass-like distributions with  $16^3$ ,  $32^3$ , and  $64^3$  particles (hereafter labeled N16, N32, and N64). To model the plane parallel radiation in URCHIN, we trace a single ray of length  $L_{\text{slab}}$  from each particle towards the surface of the slab. We calculate column densities by projecting all SPH particles



**Figure 4.** Test 2: Plane parallel radiation incident from one side onto 200 pkpc isothermal slabs of uniform density. The spectrum is that of HM01 at redshift 3 integrated between 1 and 4 Rydbergs. The slabs shown have  $n_{\text{H}} = 1.5 \times 10^{-3} \text{cm}^{-3} \times \{1.0, 0.96, 0.92, 0.88, 0.84, 0.80, 0.76, 0.72, 0.68, 0.64, 0.60\}$  (from black to red) which, at redshift 3, correspond to overdensities of  $\Delta = \{500, 480, 460, 440, 420, 400, 380, 360, 340, 320, 300\}$ . *Top panel:* equilibrium neutral fraction as a function of depth  $x(z)$  for the analytic solution (solid lines) and the URCHIN N16 resolution solution (dashed lines). Lines are labelled with the corresponding value of the over density ( $n_{\text{H}} = n_{\text{H}}^{\text{mean}} \times \Delta$ ), and the analytic neutral hydrogen column density through the slab,  $\log N_{\text{HI}}^{\text{ana}}$ . *Bottom panel:* the difference in  $N_{\text{HI}}$  between the URCHIN and analytic solutions, as a function of the analytic  $N_{\text{HI}}$ . Shown are the results when the HM01 spectrum is used (solid lines) and when the grey approximation is used (dashed lines), with different colours referring to different particle resolutions. In all cases, the URCHIN reverse ray tracing solution is within 0.1 dex of the analytic value, with the error depending weakly on particle resolution.



onto a plane and measure the mean column on a fine grid of  $2048^2$  pixels. Similar projections were used in Altay et al. (2011).

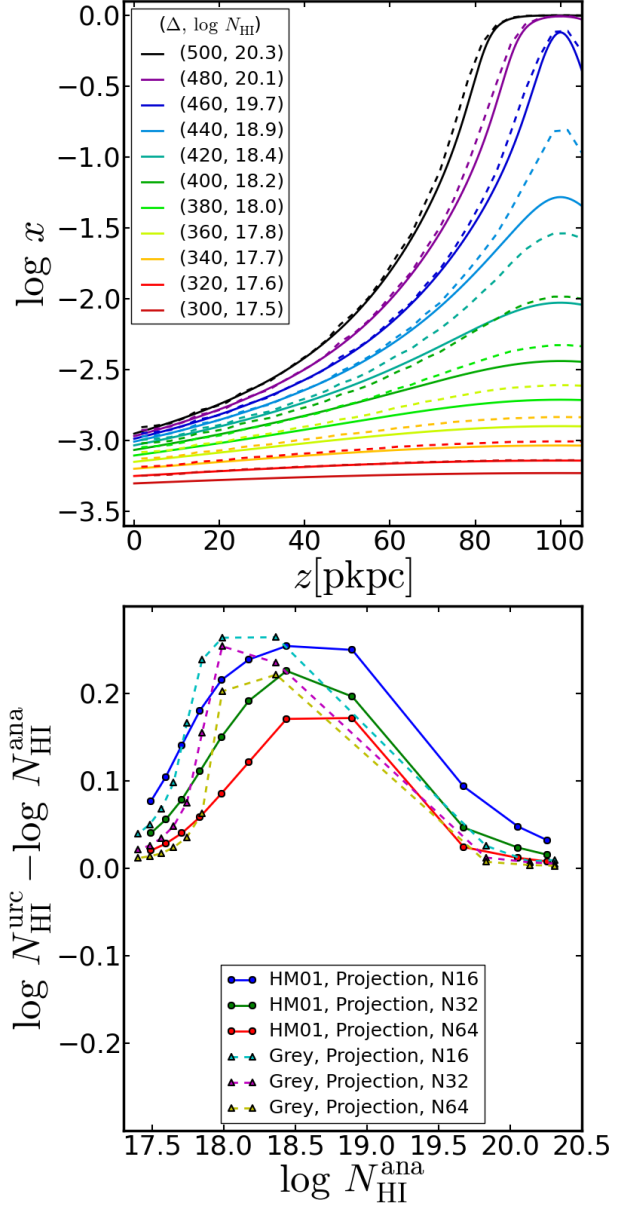
In the top panel of Figure 4, we compare  $\mathcal{NS}$  solutions (solid lines) to those produced by URCHIN (dashed lines) for slabs with different densities (different colours). The URCHIN solutions are from the lowest (N16) resolution SPH density fields. URCHIN faithfully follows the dependence of neutral fraction on depth for all models, including those where the gas becomes mostly neutral. The bottom panel of the figure quantifies the errors in neutral hydrogen column densities calculated from the SPH realizations (solid lines). Errors are below 0.05 dex at low resolution (N16, blue line), and improve with increasing resolution. The dashed lines quantify the URCHIN errors in the case of monochromatic radiation, for which the analytic solution  $\mathcal{G}$  does not require the construction of the interpolation table discussed in Section 4.1.2. The errors are slightly larger here as the ionisation front becomes very steep when the radiation is monochromatic. However, this test demonstrates that our method of splitting the optical depth into a contribution from near and far particles works well for single ray applications.

### 5.3 Test 3: Uniform Slab - Radiation Incident from Two Sides

This test is identical to Test 2, except we irradiate the slabs from both sides. To obtain the analytic solution, we modify  $\mathcal{NS}$  as follows. First we initialise the neutral fractions of all slab slices to their optically thin values, then we loop over the slices calculating the optical depth both above and below a given slice. These optical depths are used to calculate a photoionisation rate and hence a new value for the neutral fraction. We continue iterating over the slices until the neutral fraction in each slice has converged to one part in ten thousand. The solution is symmetric with respect to the centre of the slab at  $z = 100$  pkpc. To model the plane parallel radiation in URCHIN, we trace two rays from each particle, one in the  $+z$  direction and one in the  $-z$  direction.

As in the previous test, we compare the URCHIN (dashed lines) and analytic (solid lines) solutions for the neutral fraction in the top panel of Fig. 5. The goal of this test is to examine the accuracy of the *near / far* split described in section 4.1.3 when multiple rays are being used (diagrammed in panel b of Figure 2). The algorithm introduces errors in the calculation of  $\Gamma^{\text{shld}}$  because each ray is considered independently. To illustrate this point we will consider the process of calculating  $\Gamma^{\text{shld}} = 2\pi(\Gamma_+^{\text{shld}} + \Gamma_-^{\text{shld}})$  for a particle situated in the middle of a slab in which the equilibrium neutral fraction does not form a neutral core (i.e. any slab with  $\Delta < 460$ ). First the  $+z$  ray is traced and a  $\Gamma_+^{\text{far}}$  is calculated as in Eq. (12). This is then used as input to calculate  $\Gamma_+^{\text{shld}}$  as in Eq. (13) which represents the photoionisation rate at the bottom of the *near* slab. The error occurs due to the fact that the *near* slab should also be irradiated from the  $-z$  direction as all of the gas is highly ionised. This leads to an overestimate of the opacity of each of the *near* slabs and in turn to an overestimate of the neutral fraction in the slab. The errors are most severe during the transition from optically thin slabs to slabs that form a neutral core. In slabs that do form a neutral core the error is absent as at least one ray always encounters a high optical depth.

The resulting errors on the column density through the slab are shown in the bottom panel of Fig. 5. Different colours refer to different numerical resolutions, with solid lines representing the HM01 radiation field, and dashed lines the grey approximation. The URCHIN optical depth is within 0.1 dex of the analytic result for columns  $N_{\text{HI}} < 10^{18} \text{ cm}^{-2}$  or  $N_{\text{HI}} > 10^{19.5} \text{ cm}^{-2}$  for the highest resolution slab. These column density limits correspond to the cases



**Figure 5.** Test 3: Plane parallel radiation incident from two sides onto 200 pkpc isothermal slabs of uniform density. The spectrum is that of HM01 at redshift 3 integrated between 1 and 4 Rydbergs. The panels are arranged as in Figure 4. *Top panel:* for slabs with  $\Delta \leq 340$  or  $\Delta \geq 460$ , errors in the ionisation profiles are similar to Test 2. In the intermediate regime both rays are important in determining the photoionisation rate at a particle and the assumptions made in the solution  $\mathcal{G}$  are not valid (see Figure 2). *Bottom panel:* the difference in  $N_{\text{HI}}$  between the URCHIN and analytic solutions, as a function of the analytic  $N_{\text{HI}}$ . Because the importance of  $\mathcal{G}$  is reduced at higher resolutions, these errors scale more strongly with particle resolution than those in Test 2.

where the slab is either mostly ionised everywhere, or develops a neutral core. In the intermediate regime, URCHIN overestimates the neutral hydrogen column density by up to 0.15 dex at N64 resolution, and 0.25 dex at N16 resolution. The errors in the case of monochromatic radiation (dashed lines) are not significantly different, showing that the URCHIN error is due to the *near/far* split, rather than the implementation of the look-up table for polychro-

matic spectra employed in  $\mathcal{G}$  (as opposed to the analytic solution used for monochromatic spectra).

#### 5.4 Test 4: Galactic Halos - Uniform UV Background

In previous sections we applied URCHIN to very simple slab geometries for which we could calculate accurate analytic solutions. In this section we concentrate on the more realistic case of galactic halos. In particular we focus on halos extracted from the OWLS suite of cosmological galaxy formation simulations (Schaye et al. 2010). This suite consists of a reference model (REF) and more than 50 variations around that reference model which explore changes to sub-grid physics and other parameters. The REF model, which we make use of here, included a model for pressure in the numerically unresolved cold interstellar medium, star formation, the timed release of 11 chemical elements by type I and type II supernovae and AGB stars, radiative cooling due to the same elements in the presence of the HM01 ionising background, and energetic feedback from supernovae.

##### 5.4.1 Stacking Halos

To identify these halos we first calculate a friends-of-friends group catalogue using the standard algorithm of Davis et al. (1985) run on the dark matter particles, and linking baryonic particles to their nearest dark matter particle. We then identify bound sub-structures (i.e. halos) within these FoF groups using the subfind algorithm (Springel et al. 2001; Dolag et al. 2009) and associate these halos with galaxies. Our goal is to create objects that have approximate spherical symmetry – so we can calculate self-shielding analytically – yet are representative of typical galaxy density profiles encountered in simulations. To this end, we ‘stack’ haloes of similar mass. First, haloes are grouped into mass bins according to their total gas mass, with the lowest mass bin edge being  $M_{\text{gas}} = 10^{8.15} M_{\odot} h^{-1}$  (100 gas particles) and the bin width equal to 0.3 dex. We then centre all haloes in a given bin on the location of the most bound particle of that halo. The mass of each particle in a stack is then adjusted such that the sum of all particle masses in the stack is at the logarithmic center of the mass bin. Finally, the smoothing lengths and densities of all particles in a stack are recalculated. These stacks, while still containing some particle noise, are now (nearly) spherically symmetric. In Figure 6 we show images of each halo stack out to a radius of 100 pkpc. The color scale logarithmically covers the range between  $N_{\text{H}} = 10^{17} \text{cm}^{-2}$  and  $N_{\text{H}} = 10^{23} \text{cm}^{-2}$ . We show this figure mainly to give the reader a visual impression of how the halo stacks become less spherically symmetric in the higher mass bins due to the smaller number of halos in each bin.

##### 5.4.2 Radial Profiles

In Fig. (7) we plot radial density profiles of the stacked haloes. In each panel there are four black dashed lines. The lower horizontal line indicates the cosmic mean value of  $n_{\text{H}}$  at  $z = 3$ , the upper horizontal line indicates the star-formation density threshold used in REF, the vertical line indicates the gravitational softening length, and the diagonal line is an arbitrarily normalised  $1/r^2$  line to guide the eye. Both the total and neutral hydrogen number density are shown with shaded regions. The two quantities are equal at small radii causing the shaded regions to have a shape similar to the greek letter lambda in each panel.

##### 5.4.3 Total Hydrogen Radial Profiles

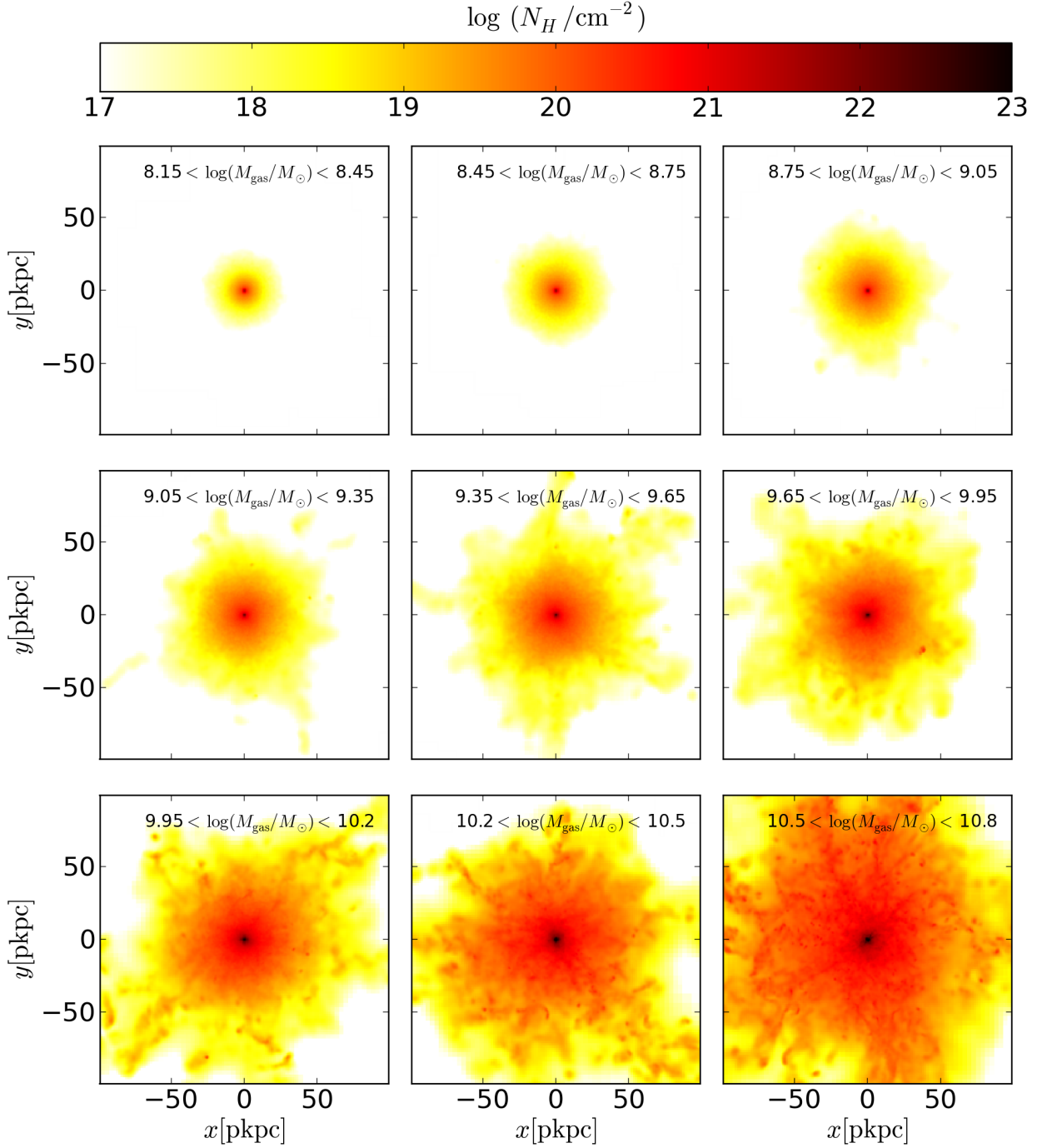
Total hydrogen density profiles were calculated from the SPH stacks and are plotted as the upper blue shaded regions. The width of the regions correspond to one sigma variations from the mean density at a given radius. This is a measure of the deviations from spherical symmetry shown in Figure 6. The shaded regions are mass weighted averages but for reference we also plot the volume weighted average, i.e. the sum of all particle masses in a radial shell divided by the volume of the shell, with a green line in the bottom left panel. The volume weighted density hugs the lower end of the mass weighted density. In an SPH distribution with no particle noise the two quantities would be equal; however, any clumping will increase the mass weighted quantity relative to the volume weighted quantity. The offset between the two at small radii indicates the level of particle noise in the spherically symmetric part of the SPH stacks while the differences at large radii are due to clumps caused by stacking a finite number of halos. Note that this noise was mostly absent in the previous tests due to the use of glass like distributions.

For each mass bin we construct two smooth analytic  $n_{\text{H}}$  profiles by fitting a polynomial through the one sigma variations described above (upper red shaded region). The differences between the two are only visible at the larger radii of the more massive bins. We use the point where the volume weighted  $n_{\text{H}}$  intersects the cosmic mean  $n_{\text{H}}$  to define a radius for each analytic profile which is why the red shaded regions are truncated at smaller radii than the blue. This provides us with a perfectly smooth and spherically symmetric approximation to our SPH stacks. In general, the density at a given radius is monotonically increasing with halo mass as is the radius of the halos. Each halo has a profile close to  $1/r^2$  at intermediate radii but becomes steeper at both smaller and larger radii.

##### 5.4.4 Neutral Hydrogen Radial Profiles

To calculate neutral hydrogen profiles in spherically symmetric gas with  $\mathcal{N}\mathcal{S}$  we did the following. Each halo is divided into  $N_{\text{sh}}$  shells. From each shell, we trace  $N_{\theta}$  rays which sample the azimuthal angle between 0 and  $\pi$  radians. The photoionisation rate in each shell is calculated by summing the contribution from each ray and is then used to calculate a new neutral fraction. We loop over all shells and iterate until the neutral fraction in each shell has converged to one part in ten thousand. We find that our results are numerically converged when using  $N_{\text{sh}} = 1600$  and  $N_{\theta} = 13$ . We use  $\mathcal{N}\mathcal{S}$  to calculate  $n_{\text{HI}}$  profiles from the  $-1\sigma$  and  $+1\sigma$   $n_{\text{H}}$  profiles, and show the results as the lower red shaded regions.

We also use URCHIN to calculate the neutral fraction of every particle in each stack and construct average  $n_{\text{HI}}$  profiles in the same way as we calculated the  $n_{\text{H}}$  SPH density profiles. These are shown as the lower blue shaded areas. When irradiated from outside, each halo develops a characteristic ionisation structure with a neutral core, a sharp ionization front, and an optically thin region. In the core,  $n_{\text{H}} = n_{\text{HI}}$ . When the density drops to  $\log n_{\text{H}} \approx -2$  the neutral core transitions to a steep ionisation front. At densities between  $\log n_{\text{H}} = -4$  and  $\log n_{\text{H}} = -5$  the ionisation profiles transition into the optically thin regime in which the neutral fraction is proportional to the density,  $x = \alpha n_{\text{H}}/\Gamma$ . In all cases, the URCHIN solution overlaps with the analytic solution. The differences between the two are largest in the ionization front but match the analytic solution in the optically thin and optically thick regimes. This is the spherical corollary to the situation in Test 3 in which the errors were largest for intermediate column density slabs. In a spherical

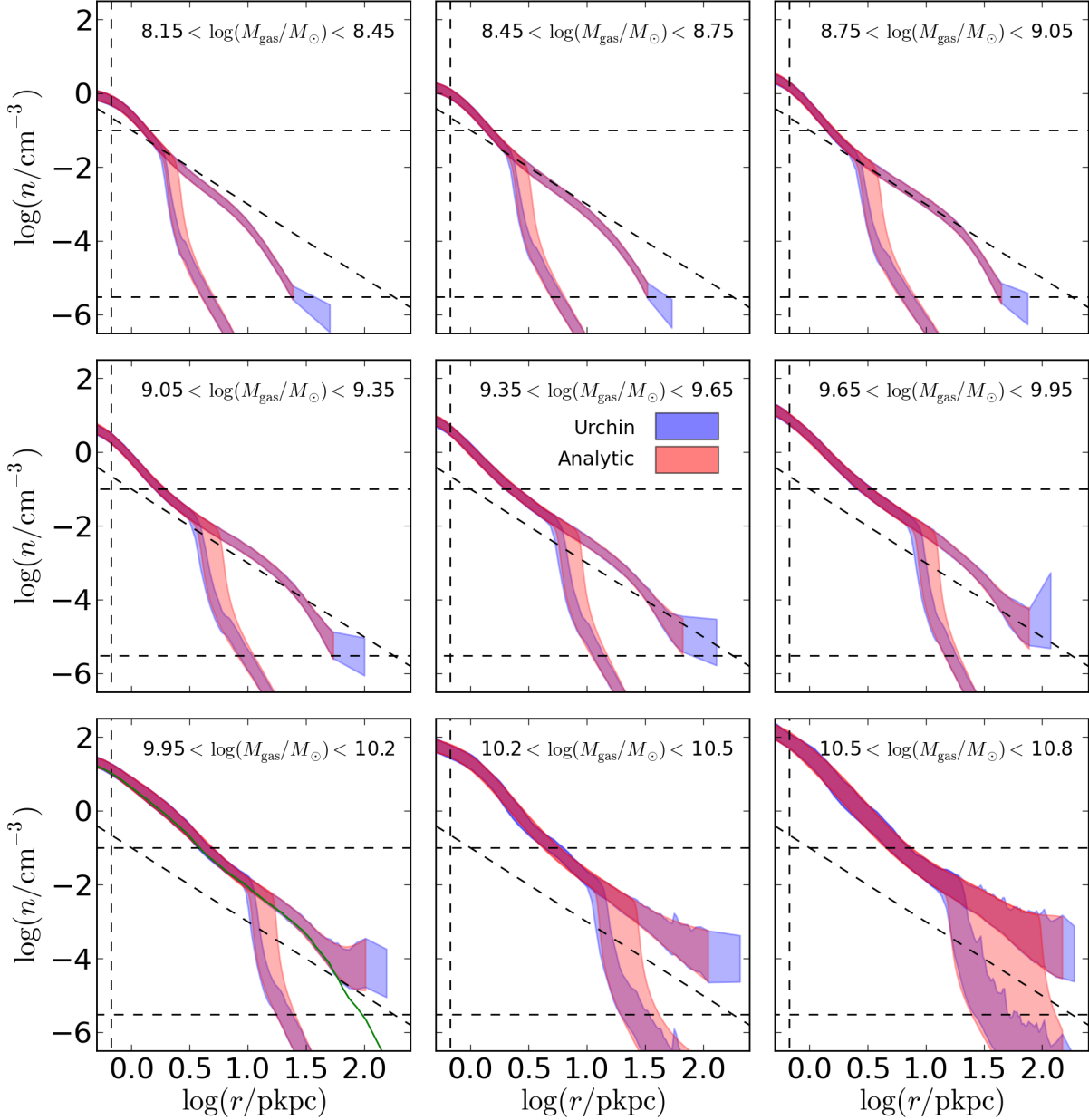


**Figure 6.** Test 4: Projected total hydrogen density for the nine halo mass bins used for constructing the halo stacks. The color scale logarithmically covers the range between  $N_{\text{H}} = 10^{17} \text{cm}^{-2}$  and  $N_{\text{H}} = 10^{23} \text{cm}^{-2}$ . Each image shows a region 200 pkpc on a side and deep enough to include every particle in the stack. These images give a sense of how the higher mass halo stacks are less spherically symmetric than the lower mass stacks, as there are fewer of them within the simulation volume to average over.

geometry, even with a neutral core, some rays will sample the problematic columns between  $10^{18} < N_{\text{HI}}/\text{cm}^{-2} < 10^{19.5} \text{cm}^{-2}$ .

#### 5.4.5 Neutral Hydrogen Column Density Profiles

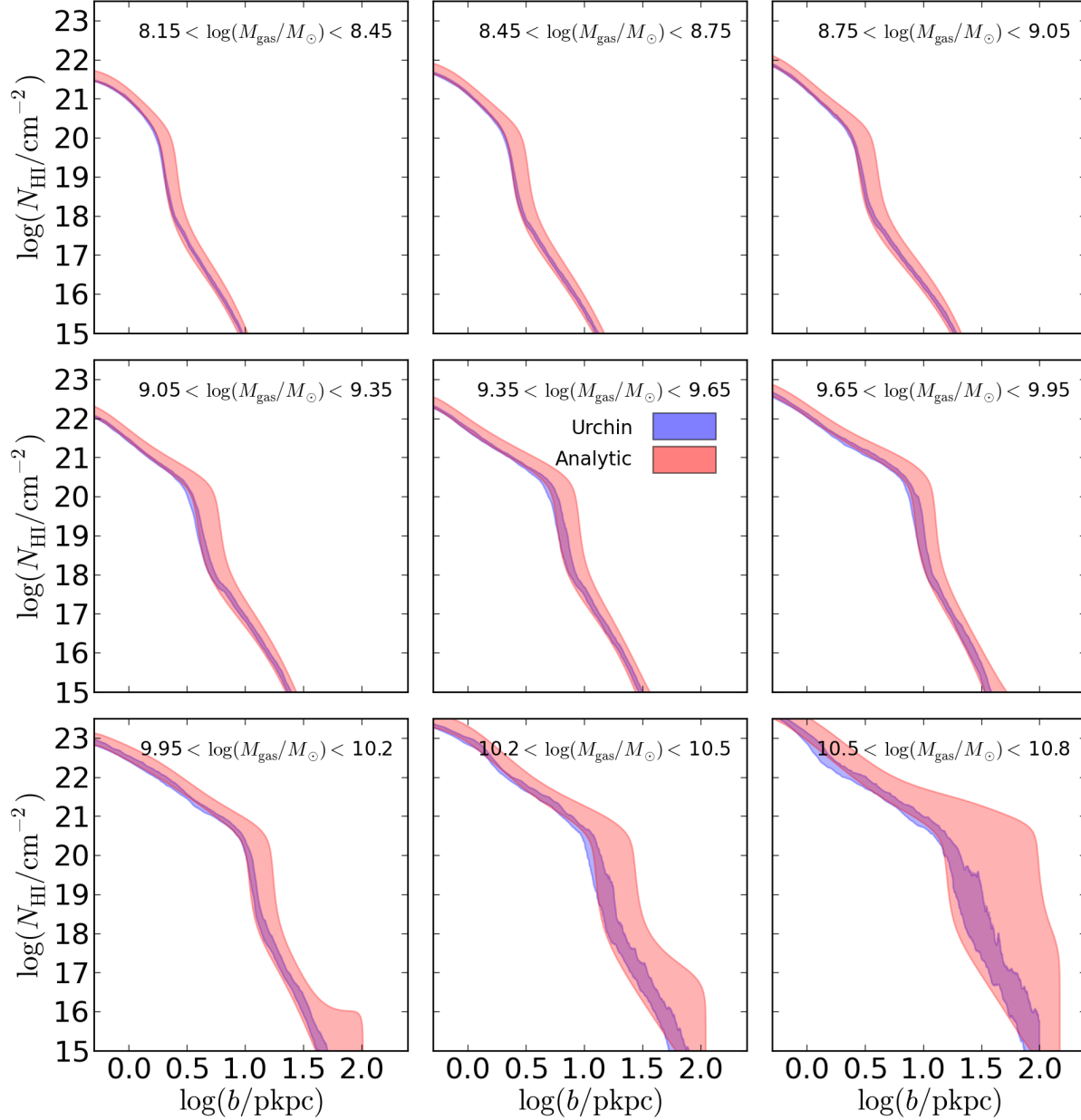
In Figure 8 we show the neutral hydrogen column density  $N_{\text{HI}}$  as a function of impact parameter for the halo stacks. To calculate  $N_{\text{HI}}$  profiles in the analytic case we simply integrate lines of sight



**Figure 7.** Test 4:  $n_{\text{H}}$  and  $n_{\text{HI}}$  radial profiles for OWLS halo stacks irradiated with a uniform UV-background. Different panels correspond to different halo gas mass ranges as indicated in each panel. The lower horizontal dashed line indicates the cosmic mean value of  $n_{\text{H}}$  at  $z = 3$ , the upper horizontal line indicates the star-formation density threshold used in the REF OWLS model, the vertical line indicates the gravitational softening length, and the diagonal line shows an arbitrarily normalised  $1/r^2$  relationship to guide the eye. Both the total ( $n_{\text{H}}$ ) and neutral ( $n_{\text{HI}}$ ) hydrogen number density are shown with shaded regions. The two quantities are equal at small radii causing the shaded regions to have a shape similar to the greek letter lambda in each panel. The SPH density fields on which URCHIN was run are shown in blue while our analytic solutions are shown in red. All SPH profiles (blue) are constructed from mass weighted averages at a given radii. The width of the shaded regions indicates one sigma variations from the median, and hence are a measure of deviations from spherical symmetry in the stacks. In the lower left panel we also indicate the volume weighted average of  $n_{\text{H}}$  with a green line. The URCHIN solution and the analytic solution always overlap. The two solutions are most different in regions where the  $n_{\text{HI}}$  profile is steepest.

through the spherically symmetric shells shown in Figure 7 and tabulate the results as a function of impact parameter  $b$ . We calculate these profiles for the  $-1\sigma$  and  $+1\sigma$  spherical mass profiles, and shade the region between the two in red. To calculate the same quantity from the SPH distributions we project all particles in a stack onto a plane and measure the column density on a fine grid of  $2048^2$  pixels. We then bin these pixels in impact parameter and

show the one sigma variation around the mean as a shaded blue region. For density profiles of the form  $n_{\text{H}} \propto r^{-n}$  with  $n > 1/2$  the dominant contribution to  $N_{\text{HI}}$  along a line of sight comes from the smallest radii. Because of this, we can associate lines of sight in Figure 8 with the innermost radius they probe. This gives rise to a corresponding region in the  $N_{\text{HI}}(b)$  plots for each of the regions described in the previous section (neutral core, ionisation



**Figure 8.** Test 4:  $N_{\text{HI}}$  impact parameter profiles for OWLS halo stacks irradiated with a uniform UV-background. Different panels correspond to different halo gas mass ranges as indicated in each panel. The SPH density fields on which URCHIN was run are shown in blue while our analytic solutions are shown in red. The SPH profiles (blue) are constructed by projecting all particles in a halo stack onto a plane. The width of the shaded regions indicates one sigma variations from the median value at a given impact parameter. The URCHIN solution and the analytic solution always overlap.

front, and optically thin outskirts). In all mass bins, the URCHIN solution overlaps with the analytic solution giving us confidence that URCHIN can be used to accurately model neutral hydrogen absorbers in a cosmological context.

## 6 DISCUSSION AND CONCLUSIONS

We have presented and described a new publicly available radiative transfer code called URCHIN. It is optimized to model the residual neutral hydrogen in the post-reionization Universe and relies on reverse ray tracing to avoid problems typically associated

with modelling the UV background. In particular, our implementation allows for the preservation of symmetry in the input radiation field, a completely uniform sampling of gas resolution elements, a high fidelity sampling of spectral features, and some optimisations not possible in forward ray tracing schemes.

We have validated URCHIN in four sets of tests that range from comparison to analytic solutions in simple slab geometries to solving for the ionization structure in gaseous galactic halos. The errors discussed in §5 have two root causes. One is particle resolution, and the other is the multi-valued relationship between points in space and resolution elements in SPH. The only solution to the first prob-

lem is to run simulations with more particles. The second problem prompted us to introduce the *near / far* split and the analytic slab solution  $\mathcal{G}$  discussed in §4. This solution solves the multi-valued problem but is restrictive in its use of plane parallel geometry and its treatment of each ray / *near* slab pair independently (i.e. without considering the other pairs). It is possible that an approach that parameterises a spherical analytic solution using some average of the *near* slab optical depths and some average of the incoming radiation from all rays as parameters would be more successful. In addition, the analytic solution presented in Eq. A22 considers uniform density slabs although there is a density derivative in the fully general solution. Estimating  $dn_{\text{H}}/dr$  along each ray and using this as an extra parameter in  $\mathcal{G}$  could also prove useful in improving the solution.

In the current implementation, we have shown that solutions generated by URCHIN reproduce analytic solutions in the case of density fields approximating gaseous galactic halos. We view URCHIN as a first step in a series of incremental improvements to the standard treatment of the UV background. Future steps will include helium as in Altay et al. (2008), proximity zones from internal point sources, an equilibrium heating/cooling model, and non-equilibrium corrections for ionisation and heating.

## ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation under Grant No. NSF PHY11-25915. The calculations for this paper were performed on the ICC Cosmology Machine, which is part of the DiRAC Facility jointly funded by STFC, the Large Facilities Capital Fund of BIS, and Durham University.

## REFERENCES

- Abel T., Norman M. L., Madau P., 1999, *ApJ*, 523, 66
- Altay G., Croft R. A. C., Pelupessy I., 2008, *MNRAS*, 386, 1931
- Altay G., Theuns T., Schaye J., Crighton N. H. M., Dalla Vecchia C., 2011, *ApJ*, 737, L37
- Aubert D., Teyssier R., 2008, *MNRAS*, 387, 295
- Battisti A. J., Meiring J. D., Tripp T. M., et al., 2012, *ApJ*, 744, 93
- Becker G. D., Rauch M., Sargent W. L. W., 2007, *ApJ*, 662, 72
- Bi H., Davidsen A. F., 1997, *ApJ*, 479, 523
- Cen R., 2012, *ApJ*, 748, 121
- Cen R., Miralda-Escudé J., Ostriker J. P., Rauch M., 1994, *ApJ*, 437, L9
- Croft R. A. C., 2004, *ApJ*, 610, 642
- Darling J., Macdonald E. P., Haynes M. P., Giovanelli R., 2011, *ApJ*, 742, 60
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371
- Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, 399, 497
- Duffy A. R., Kay S. T., Battye R. A., Booth C. M., Dalla Vecchia C., Schaye J., 2012, *MNRAS*, 420, 2799
- Efstathiou G., 1992, *MNRAS*, 256, 43P
- Erkal D., Gnedin N. Y., Kravtsov A. V., 2012, *ArXiv e-prints*
- Fan X., Carilli C. L., Keating B., 2006, *ARA&A*, 44, 415
- Finlator K., Özel F., Davé R., 2009, *MNRAS*, 393, 1090
- Fumagalli M., Prochaska J. X., Kasen D., Dekel A., Ceverino D., Primack J. R., 2011, *MNRAS*, 418, 1796
- Gingold R. A., Monaghan J. J., 1977, *MNRAS*, 181, 375
- Gnedin N. Y., Abel T., 2001, *New A*, 6, 437
- Górski K. M., Hivon E., Banday A. J., et al., 2005, *ApJ*, 622, 759
- Gunn J. E., Peterson B. A., 1965, *ApJ*, 142, 1633
- Gupta N., Srianand R., Bowen D. V., York D. G., Wadadekar Y., 2010, *MNRAS*, 408, 849
- Haardt F., Madau P., 2001, in *Clusters of Galaxies and the High Redshift Universe Observed in X-rays*, edited by D. M. Neumann & J. T. V. Tran
- Haardt F., Madau P., 2012, *ApJ*, 746, 125
- Haehnelt M. G., Steinmetz M., Rauch M., 1998, *ApJ*, 495, 647
- Hernquist L., Katz N., Weinberg D. H., Miralda-Escudé J., 1996, *ApJ*, 457, L51
- Iliev I. T., Ciardi B., Alvarez M. A., et al., 2006, *MNRAS*, 371, 1057
- Iliev I. T., Whalen D., Mellema G., et al., 2009, *MNRAS*, 400, 1283
- Katz N., Weinberg D. H., Hernquist L., Miralda-Escudé J., 1996, *ApJ*, 457, L57
- Köhler K., Gnedin N. Y., 2007, *ApJ*, 655, 685
- Lucy L. B., 1977, *AJ*, 82, 1013
- Martin A. M., Papastergis E., Giovanelli R., Haynes M. P., Springob C. M., Stierwalt S., 2010, *ApJ*, 723, 1359
- Maselli A., Ferrara A., 2005, *MNRAS*, 364, 1429
- Maselli A., Ferrara A., Ciardi B., 2003, *MNRAS*, 345, 379
- McQuinn M., Oh S. P., Faucher-Giguère C.-A., 2011, *ApJ*, 743, 82
- Meiksin A. A., 2009, *Reviews of Modern Physics*, 81, 1405
- Mellema G., Iliev I. T., Alvarez M. A., Shapiro P. R., 2006, *New A*, 11, 374
- Mirocha J., Skory S., Burns J. O., Wise J. H., 2012, *ArXiv e-prints*
- Monaghan J. J., Lattanzio J. C., 1985, *A&A*, 149, 135
- Mortlock D. J., Warren S. J., Venemans B. P., et al., 2011, *Nature*, 474, 616
- Myers A. D., Palanque-Delabrouille N., Schlegel D. J., et al., 2012, in *American Astronomical Society Meeting Abstracts*, vol. 219 of *American Astronomical Society Meeting Abstracts*, 335.12
- Nakamoto T., Umemura M., Susa H., 2001, *MNRAS*, 321, 593
- Noterdaeme P., Petitjean P., Carithers W. C., et al., 2012, *A&A*, 547, L1
- Okamoto T., Gao L., Theuns T., 2008, *MNRAS*, 390, 920
- O’Meara J. M., Prochaska J. X., Worseck G., Chen H.-W., Madau P., 2012, *ArXiv e-prints*
- Osterbrock D. E., 1989, *Astrophysics of gaseous nebulae and active galactic nuclei*, University Science Books
- Paardekooper J.-P., Kruip C. J. H., Icke V., 2010, *A&A*, 515, A79
- Pawlik A. H., Schaye J., 2008, *MNRAS*, 389, 651
- Pawlik A. H., Schaye J., 2011, *MNRAS*, 412, 1943
- Petkova M., Springel V., 2009, *MNRAS*, 396, 1383
- Pontzen A., Governato F., Pettini M., et al., 2008, *MNRAS*, 390, 1349
- Portegies Zwart S., McMillan S., Harfst S., et al., 2009, *New A*, 14, 369
- Prochaska J. X., Herbert-Fort S., 2004, *PASP*, 116, 622
- Prochaska J. X., Worseck G., O’Meara J. M., 2009, *ApJ*, 705, L113
- Rahmati A., Schaye J., Pawlik A. H., Raičević M., 2013, *ArXiv e-prints*
- Razoumov A. O., Cardall C. Y., 2005, *MNRAS*, 362, 1413
- Razoumov A. O., Norman M. L., Prochaska J. X., Wolfe A. M., 2006, *ApJ*, 645, 55

- Rijkhorst E.-J., Plewa T., Dubey A., Mellema G., 2006, *A&A*, 452, 907
- Rybicki G. B., Lightman A. P., 1986, *Radiative Processes in Astrophysics*
- Schaye J., 2001, *ApJ*, 559, 507
- Schaye J., 2006, *ApJ*, 643, 59
- Schaye J., Dalla Vecchia C., Booth C. M., et al., 2010, *MNRAS*, 402, 1536
- Schlegel D. J., Blanton M., Eisenstein D., et al., 2007, in *American Astronomical Society Meeting Abstracts*, vol. 38 of *Bulletin of the American Astronomical Society*, 132.29
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726
- Susa H., 2006, *PASJ*, 58, 445
- Theuns T., Leonard A., Efstathiou G., 1998a, *MNRAS*, 297, L49
- Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998b, *MNRAS*, 301, 478
- van de Voort F., Schaye J., Altay G., Theuns T., 2012, *MNRAS*, 2664
- Verner D. A., Ferland G. J., Korista K. T., Yakovlev D. G., 1996, *ApJ*, 465, 487
- Whalen D., Norman M. L., 2006, *ApJS*, 162, 281
- Wise J. H., Abel T., 2011, *MNRAS*, 414, 3458
- Yajima H., Choi J.-H., Nagamine K., 2011, *ArXiv e-prints*

## APPENDIX A: ANALYTIC HYDROGEN IONIZATION SOLUTIONS

### A1 Photoionization Cross Section

The photoionization cross section for hydrogenic atoms in the ground state can be expressed analytically as,

$$\sigma = \frac{\sigma_{\text{th}}}{Z^2} \left( \frac{\nu_{\text{th}}}{\nu} \right)^4 \frac{\exp \{4 - [(4 \tan^{-1} \epsilon)/\epsilon]\}}{1 - \exp(-2\pi/\epsilon)}$$

$$\epsilon = \sqrt{\frac{\nu}{\nu_{\text{th}}}} - 1, \quad \sigma_{\text{th}} = \frac{2^9 \pi}{3e^4} \alpha_{\text{fs}} \pi a_0^2, \quad (\text{A1})$$

where  $\alpha_{\text{fs}}$  is the fine structure constant,  $a_0$  is the Bohr radius, and  $Z$  is the atomic number of the atom. Two fitting formula are commonly used in the literature. The first is a power law form, usually accompanied by a citation to Osterbrock (1989),

$$\sigma = \sigma_{\text{th}} \left( \frac{\nu}{\nu_{\text{th}}} \right)^{-3}. \quad (\text{A2})$$

This simple form can be useful for analytic treatments, but is not as accurate as the fit due to Verner et al. (1996),

$$\sigma = \sigma_0 (x - 1)^2 x^{0.5P - 5.5} \left( 1 + \sqrt{x/x_a} \right)$$

$$x \equiv \frac{h\nu}{E_0}$$

$$E_0 = 0.4298 \text{ eV}, \quad \sigma_0 = 5.475 \times 10^4 \text{ Mb}$$

$$x_a = 32.88, \quad P = 2.963. \quad (\text{A3})$$

We explore the errors these approximations introduce into column density calculations in the next section.

### A2 Power-law approximations to the Haardt & Madau 2001 spectrum

The left panel in Figure A1 compares the shape of the Haardt & Madau (2001) (HM01) spectrum to power-law approximations of

the form  $J(\nu) = J_0 (\nu/\nu_{\text{th}})^{-\alpha}$ , for  $\alpha = 0, 1, 2$ . The HM01 spectrum has features due to re-emission, such as the Helium Lyman- $\alpha$  emission at 3 Rydberg. The effect of using such approximate spectra as opposed to the full HM01 spectrum is illustrated in the middle panel. We calculated the equilibrium neutral fraction as function of depth,  $x_{\text{eq}}(z)$ , for plane-parallel radiation entering a slab with uniform hydrogen density  $n_{\text{H}} = 1.5 \times 10^{-3} \text{ cm}^{-3}$  (500 times the cosmic mean  $n_{\text{H}}$  at  $z = 3$ ) and  $T = 10^4 \text{ K}$ . Power-law approximations to HM01 work reasonably well, as long as the spectra are normalised to give the same number density of ionising photons as the original HM01 spectrum (dashed lines). Normalising the power-law spectra to the same photoionisation rate does not work well (solid lines). The right panel in Figure A1 compares the differences in neutral columns at  $z = 0.2 \text{ pMpc}$ , between the full HM01 spectrum and these powerlaw approximations. The largest error occurs when the slab fails to become fully neutral, as in the flat spectrum case ( $\alpha = 0$ ). The monochromatic grey approximation used in Section 5.1 (pink bar) overestimates the central HI column by  $\leq 0.04 \text{ dex}$ . The bottom two bars show the effect of approximating the photoionisation cross-section by a simple power law versus using the more accurate fit from Verner et al. (1996), and using the HM01 spectrum up to 10 Rydberg, as opposed to 4 Rydberg as we have done up to now.

### A3 Time dependent solution of the neutral fraction

The rate of change of the hydrogen neutral fraction  $x \equiv n_{\text{HI}}/n_{\text{H}}$  is determined by the rate of photoionisation  $\Gamma$ , collisional ionisation  $\gamma$ , and recombination  $\alpha$  as well as the number density of free electrons  $n_e$ .

$$\frac{dx}{dt} = -(\Gamma + \gamma n_e)x + \alpha n_e(1 - x) \quad (\text{A4})$$

If we decompose the free electron number density into a contribution from ionised hydrogen and a contribution  $yn_{\text{H}}$  from all heavier elements, we can write:

$$n_e = (1 - x + y)n_{\text{H}} \quad (\text{A5})$$

Substituting into the previous equation yields:

$$\frac{dx}{dt} = -[\Gamma + \gamma(1 - x + y)n_{\text{H}}]x + \alpha(1 - x + y)n_{\text{H}}(1 - x) \quad (\text{A6})$$

Grouping terms in powers of  $x$ , we can write  $dx/dt$  in the form of a Riccati equation:

$$\frac{dx}{dt} = Rx^2 + Qx + P \quad (\text{A7})$$

$$R \equiv (\gamma + \alpha)n_{\text{H}}$$

$$Q \equiv -[\Gamma + (\gamma + 2\alpha)n_{\text{H}} + (\gamma + \alpha)n_{\text{H}}y]$$

$$= -[\Gamma + \alpha n_{\text{H}} + R(1 + y)]$$

$$P \equiv \alpha n_{\text{H}}(1 + y). \quad (\text{A8})$$

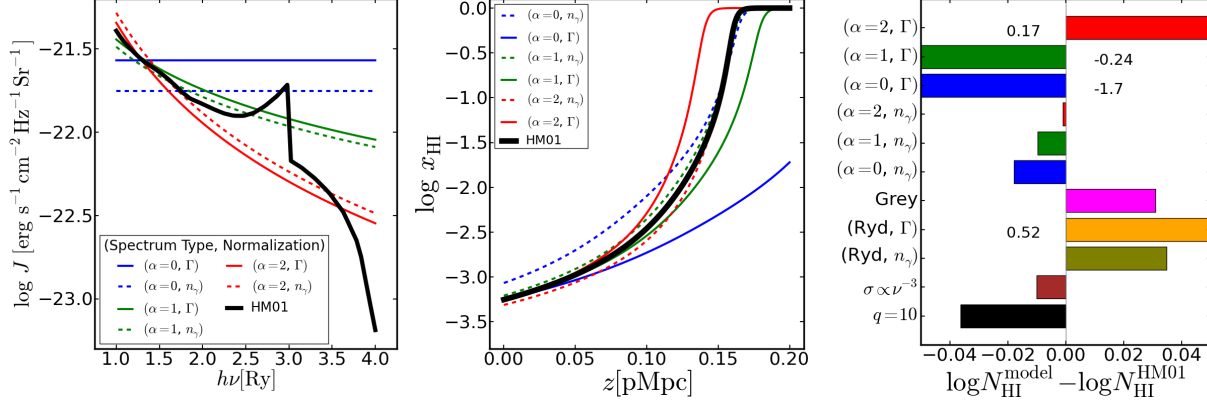
The roots of the quadratic term are

$$x_- = \frac{-Q - (Q^2 - 4PR)^{1/2}}{2R} \quad (\text{A9})$$

$$x_+ = \frac{-Q + (Q^2 - 4PR)^{1/2}}{2R}. \quad (\text{A10})$$

To determine which of these roots is the physical equilibrium solution we consider the case of pure hydrogen ( $y = 0$ ) in the absence





**Figure A1.** *Left panel:* The  $z = 3$  Haardt & Madau (2001) UV-background (HM01) spectrum compared to three different power-law spectra,  $J(\nu) = J_0 (\nu/\nu_{\text{th}})^{-\alpha}$ , for  $\alpha = 0, 1, 2$  shown blue, green and red, respectively. For the solid lines,  $J_0$  is chosen such that the power-law spectrum has the same photoionisation rate as the HM01 spectrum, for the dashed lines  $J_0$  is chosen such that they have the same number density of ionising photons. The spike around 3 Rydbergs is due to Helium Lyman- $\alpha$  emission. *Middle panel:* equilibrium neutral fraction,  $x_{\text{eq}}(z)$  as a function of depth  $z$  into a slab with uniform hydrogen density  $n_{\text{H}} = 1.5 \times 10^{-3} \text{cm}^{-3}$  (500 times the cosmic mean  $n_{\text{H}}$  at  $z = 3$ ) and  $T = 10^4$  K. The thick black line is for the full HM01 spectrum, blue, green and red are the corresponding power-law approximations. Such power-law models work relatively well, provided the amplitude  $J_0$  of the spectrum is chosen such that the spectra have the same number density of ionising photons (dashed lines). *Right panel:* difference between the total  $N_{\text{HI}}$  for these spectral approximations and that in the HM01 model. For the four bars that go off the plot, we have indicated the horizontal value in the panel. The error in  $N_{\text{HI}}$  remains below 0.05 dex for all cases except those in which the approximating spectra are normalized to have the same optically thin photoionization rate in which case the errors can be much larger. Of course the neutral fraction in the optically thin case is incorrect unless spectra are normalised to have the same optically thin photoionization rate. For reference, we also show as a brown bar differences in the hydrogen columns for the HM01 spectrum due to using a simple power law for the photoionisation cross section, as opposed to using the fit from Verner et al. (1996), see Eq.(A3). Finally the black bar shows the effect of truncating the HM01 spectrum at 10 Rydberg, as opposed to 4 Rydberg.

of radiation ( $\Gamma = 0$ ). In this case,

$$P = \alpha n_{\text{H}} \quad (\text{A11})$$

$$R = (\alpha + \gamma) n_{\text{H}} \quad (\text{A12})$$

$$-Q = R + P \quad (\text{A13})$$

$$Q^2 - 4PR = (R - P)^2 \quad (\text{A14})$$

$$x_- = \frac{(R + P) - (R - P)}{2R} = \frac{\alpha}{\alpha + \gamma} = x_{\text{eq}} \quad (\text{A15})$$

$$x_+ = \frac{(R + P) + (R - P)}{2R} = 1 \quad (\text{A16})$$

The collisional ionisation and recombination rates depend on temperature  $T$  (see for example the fits in Theuns et al. 1998b) which, in general, will change as  $x$  changes. However in the case of constant  $T$ ,  $n_{\text{H}}$ ,  $y$ , and  $\Gamma$ , the coefficients  $P$ ,  $Q$  and  $R$  are constants as well. We can rewrite the derivative using Vieta's formula as,

$$\frac{dx}{dt} = R(x - x_{\text{eq}})(x - x_+) \quad (\text{A17})$$

and solve for the time-dependent solution by separation of variables

$$\frac{dx}{R(x - x_{\text{eq}})(x - x_+)} = dt \quad (\text{A18})$$

$$\begin{aligned} x(t) &= x_{\text{eq}} + (x_0 - x_{\text{eq}}) \frac{(x_+ - x_{\text{eq}}) F}{(x_+ - x_0) + (x_0 - x_{\text{eq}}) F} \\ F(t) &\equiv \exp \left[ \frac{-(x_+ - x_{\text{eq}}) t}{t_{\text{rc}}} \right] \\ t_{\text{rc}} &\equiv \frac{1}{(\alpha + \gamma) n_{\text{H}}} \end{aligned} \quad (\text{A19})$$

In a gas composed only of hydrogen and helium,  $y$  is bound between 0 and  $(1 - X)/(2X)$  where  $X$  is the hydrogen mass fraction of the gas. The solution  $x(t)$  is fully determined once values for  $T$ ,

$n_{\text{H}}$ ,  $y$ , and  $\Gamma$  are specified. In addition, all of the dependence on the ionisation state of elements other than hydrogen is contained in the variable  $y$ . For all of the tests performed in this paper we set  $y = 0$ .

#### A4 Neutral fraction in a plane parallel slab

In the case of plane parallel monochromatic radiation with a photon flux  $F$  incident on a semi-infinite slab of hydrogen gas with constant density and temperature, the ionisation structure  $x(z)$  can be calculated analytically. We orient our coordinate system such that the surface of the slab is coincident with the  $x - y$  plane and the positive  $z$ -axis extends into the slab. The photoionisation rate at  $z$  is  $\Gamma(z) = F \sigma e^{-\tau(z)} = \Gamma_0 e^{-\tau(z)}$ , where  $\Gamma_0$  is the photoionisation rate at  $z = 0$  and  $\sigma$  is the photoionisation cross-section for the monochromatic radiation. In equilibrium, Eq. (A6) with  $y = 0$  implies:

$$\frac{\Gamma}{\Gamma_0} = \text{Exp} \left[ - \int_0^z n_{\text{H}} x \sigma ds \right] = \frac{n_{\text{H}} \mathcal{X}}{\Gamma_0 x} \quad (\text{A20})$$

where we have defined  $\mathcal{X} \equiv \alpha(1 - x)^2 - \gamma(1 - x)x$ . Taking the logarithm of the last two terms and differentiating with respect to  $z$  gives:

$$n_{\text{H}} \sigma = -\frac{1}{x} \left\{ \frac{1}{n_{\text{H}}} \frac{dn_{\text{H}}}{dz} + \frac{d \ln \mathcal{X}}{dz} - \frac{d \ln \Gamma_0 x}{dz} \right\} \quad (\text{A21})$$

Setting the derivative of  $n_{\text{H}}$  to zero and integrating both sides over the interval  $[0, z]$  yields the depth  $N_{\text{H}} \sigma = z n_{\text{H}} \sigma$  at which the neutral fraction is  $x$ . We can write this inverse solution in terms of the equilibrium neutral fraction in the absence of radiation (i.e. in collisional ionisation equilibrium),  $x_{\text{ce}} \equiv \alpha/(\alpha + \gamma) = x(z \rightarrow \infty)$ , and the value of  $x$  at the surface of the slab  $x(z = 0) = x_0$ . The

solution  $z(x)$  is then:

$$zn_{\text{H}}\sigma = \left(\frac{1}{x_0} - \frac{1}{x}\right) + \ln\left[\frac{x(1-x_0)}{x_0(1-x)}\right] + \frac{1}{x_{\text{ce}}} \ln\left[\frac{x(x_{\text{ce}}-x_0)}{x_0(x_{\text{ce}}-x)}\right] \quad (\text{A22})$$

This allows  $x(z)$  to be mapped out. Once  $x$  is known, a photoionisation rate can be determined using Eq. A20. We use this solution to verify our numerical solver ( $\mathcal{NS}$  in the text) and then use  $\mathcal{NS}$  in more general (non-monochromatic) cases to verify URCHIN. In addition, Eq. (A22) forms the basis of the solution  $\mathcal{G}$ .